

Proceedings of:

Advances in Surface Metrology

1997 Spring Topical Meeting

June 2-5, 1997

St. John's College

Annapolis, Maryland

The American Society for Precision Engineering (ASPE) is a multidisciplinary professional and technical society concerned with research and development, design, manufacture and measurement of high accuracy components and systems. ASPE activities encompass relevant aspects of mechanical, electronic, optical and production engineering, physics, chemistry, and computer and materials science. Membership is open to anyone interested in any aspect of precision engineering.

Founded in 1986, ASPE provides a new focus for a diverse but important community. Other professional organizations have covered aspects of precision engineering, always as a sideline to their principal goals. ASPE is based on the core of generic concepts necessary to achieve precision in any application; independent of discipline, ASPE intends to be the focus for precision technology — and to represent all facets from research to application.

ASPE — The American Society for Precision Engineering
401 Oberlin Road, Suite 108
P.O. Box 10826
Raleigh, NC 27605-0826
Phone (919) 839-8444 Fax (919) 839-8039

Preface

This book comprises the proceedings of the 1997 Spring Topical Meeting. The contributions reflect the authors' opinions and are published as presented to ASPE, without change. Their inclusion in this publication does not necessarily constitute endorsement by the ASPE, or its editorial staff.

Individual readers of this book and nonprofit libraries acting for them are permitted to make fair use of the material contained in this book to be used in teaching or research; provided they copy said paper in total, as written by the author. Permission is granted to quote excerpts from the scientific or technical works in this book with proper acknowledgment of the source to include: the author's name, the book name, and the ASPE tag line. Reproduction of figures and tables is likewise permitted from the contents of this book, provided that the same acknowledgment of the same source and content is printed.

Any other type reproduction, systematic copying or multiple reproduction of any material in this book is expressly prohibited except with the written permission of the ASPE editorial staff, addressed to the Society's headquarters.

ISBN 1-887706-17-8

1997

American Society for Precision Engineering
401 Oberlin Road, Suite 108
Raleigh, NC 27605-0826

Printed in the United States of America

1997 Spring Topical Meeting

Contents

<i>Preface</i>	<i>ii</i>
<i>Contents</i>	<i>iii</i>
<i>Program</i>	<i>iv</i>
<i>Foreword</i>	<i>v</i>
<i>Technical Paper Index</i>	<i>vi-vii</i>
<i>Technical Papers</i>	<i>3-87</i>
<i>Author Index</i>	<i>89</i>

1997 Spring Topical Meeting

Meeting Schedule

Monday, June 2		Tuesday, June 3	Wednesday, June 4	Thursday, June 5
	7:30	REGISTRATION	CONTINENTAL BREAKFAST	
	8:30	SENSOR TECHNOLOGY	NOVEL MEASURING INSTRUMENTS	INTERESTING APPLICATIONS
	10:00			
		BREAK		BREAK
	10:30	SIGNAL PROCESSING I	CALIBRATION AND STANDARDS	SUMMARY
	12:00			
		LUNCH		CLOSE OF CONFERENCE
	1:30	SIGNAL PROCESSING II	OPEN FOR DISCUSSIONS	
	3:30			
		OPEN FOR DISCUSSIONS		
5:00				
ON-SITE REGISTRATION	6:30	DINNER		
	7:30	ROUND TABLE: INTEGRATION OF SURFACE METROLOGY IN MANUFACTURING	ROUND TABLE: WHAT WILL THE SURFACE LOOK LIKE IN THE FUTURE?	
	9:30			

Foreword

In recent years there has been a proliferation of instruments and technology used for surface measurements. These instruments are being used to measure surfaces for applications to optical, semiconductor and metal cutting industries. With a constant push towards lower contact forces and higher measurement speeds, it is apparent that there is a corresponding need for a better understanding of both the physics of these processes and methods for parametric or visual data representation. The focus of this meeting will be on the different technologies, measurement processes, data processing, and calibration techniques used for surface measurements.

The key objective is to bring together specialists from industry, government and academia to provide a forum for the exchange of ideas about the state-of-the-art in surface measurement covering a broad range of technologies and to highlight potential areas for future development.

***Meeting Co-Chairmen
of the 1997 Spring Topical Meeting***

*Jay Raja and Stuart T. Smith
University of North Carolina at Charlotte*

Technical Paper Index

Session I

Sensor Technology

Tuesday, June 3, 8:30 a.m. - 10:00 a.m.

1. **Opening Remarks**
Stuart T. Smith and Jay Raja (University of North Carolina-Charlotte)
2. **X-Ray Surface Metrology ***
D. K. Bowen (Bede Scientific Incorporated)
3. **Surface Characterization Using a Resonator Based Profilometer (Invited Paper)**
S. M. Harb (Coventry University), and S. T. Smith (University of North Carolina-Charlotte) page 3
4. **US TAG for ISO TC213 Issues and Developments of Standards ***
To be announced

Session II

Signal Processing I

Tuesday, June 3, 10:30 a.m. - 12:00 p.m.

1. **Filtering of Surface Profiles, Past, Present and Future**
X. Liu and J. Raja (University of North Carolina-Charlotte) page 10
2. **Estimation of Arithmetic Mean Wavelength for the Areal Topographic Data**
K. Yanagi, Y. Kobayashi and S. Hara (Nagaoka University of Technology) page 15
3. **Evaluation of Functional Features of Orthopaedic Joint Prostheses Surfaces Using Wavelets**
X. Q. Jiang, L. Blunt and K. J. Stout (University of Birmingham) page 20
4. **A New Reference Line to Surface Roughness Evaluation**
Chen, Q., Jiang, X. Q., Yang, S. and Li, Z. (Huazhong University of Science & Technology) page 26

Session III

Signal Processing II

Tuesday, June 3, 1:30 p.m. - 3:30 p.m.

1. **An Investigation of Filtering of 3D Surface Texture Measurements Using Scale-sensitive Fractal Analysis and PSD**
D. K. Cohen (Michigan Metrology), and C. A. Brown, W. A. Johnsen, P. Hoch (Worcester Polytechnic Institute) page 30
2. **A Fractal-based Comparison of Surface Profiling Instrumentation**
M. Malburg (Cummins Engine Company, Inc.) page 36
3. **A New Comprehensive Measuring and Analyzing System for Form and Surface Topography**
Xiao, S., Jiang, X.Q., Xie, T. and Li, Z. (Huazhong University of Science & Technology) page 42
4. **Study on Polishing Mechanism of Magnetic Disc Substrate**
X. M. Liu, G. Q. Cai (Northeastern University), and X. L. Feng (HEN LAN Electronic Equipment, Inc.) page 46

Tuesday, June 3, 7:30 p.m. - 9:30 p.m.

Round Table Discussion: Integration of Surface Metrology in Manufacturing

Session IV

Novel Measuring Instruments

Wednesday, June 4, 8:30 a.m. - 10:00 a.m.

1. **Phase Grating Interferometer Gauge**
D. Mansfield and I. K. Buehring (Taylor Hobson Ltd.) page 51
2. **Wide-field Scanning White Light Interferometry of Rough Surfaces**
J. Roth and P. de Groot (Zygo Corporation) page 57
3. **Expanded Area Interferometric Surface Profiling with High Spatial Resolution**
P. J. Caber, J. C. Wyant and J. Schmit (WYKO Corporation) page 61
4. **Rapid Data Acquisition and Analysis for Part Profiling**
R. Jennings (ICAMP, Inc.) page 65
3. **A Novel Metrological Low Force Probe ***
V. Badami (University of North Carolina-Charlotte)

Session V

Calibration and Standards

Wednesday, June 4, 10:30 a.m. - 12:00 p.m.

1. **Calibration Problems in Surface Metrology**
T. V. Vorburger, R. G. Dixon, J. F. Song, T. B. Renegar, J. Fu (Nat'l Institute of Standards and Technology),
V. W. Tsai (University of Maryland), and T. Doi (Nat'l Research Laboratory of Metrology) page 69
2. **Measurements of Pitch, Height, and Width Artifacts with the NIST Calibrated Atomic Force Microscope**
R. G. Dixon, T. V. Vorburger, J. Fu, R. Köning (Nat'l Institute of Standards and Technology), and
V. W. Tsai, E. D. Williams, X. S. Wang (University of Maryland) page 70
3. **Accurate Force Measurements for Miniature Mechanical Systems: A Review of Progress ***
L. P. Howard (Nat'l Institute of Standards and Technology)

Wednesday, June 4, 7:30 p.m. - 9:30 p.m.

Round Table Discussion: What Will the Surface Look Like in the Future?

Session VI

Interesting Applications

Thursday, June 5, 8:30 a.m. - 10:00 a.m.

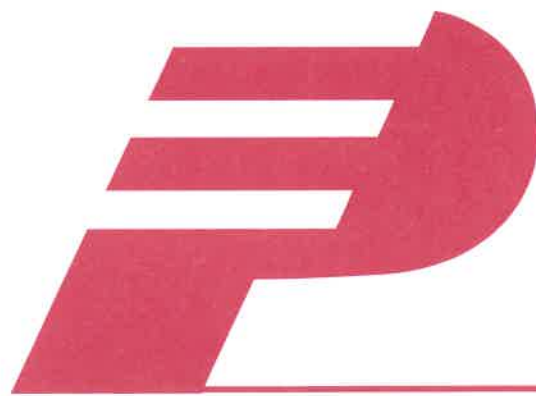
1. **AFM/IFM Nanometer Scale Mechanical Property Measurements ***
K. Jaraush and P. Russell (North Carolina State University)
2. **Atomic Force Microscopy for Profiler Stylus Quality Control**
H. H. Gatzert, J. C. Maetzig and H. Siekmann (Hanover University) page 76
3. **Atomic Force Microscope Applications in Surface Metrology ***
C. Mooney (Park Scientific Instruments)
4. **Instrument for On-line Monitoring of Surface Roughness of Machined Surfaces**
J. G. Valliant (Optical Dimensions), and J. M. Bennett (Consultant) page 82

Summary

Thursday, June 5, 10:30 a.m. - 12:00 p.m.

* Abstract not available at press time

Technical Papers



A S P E

Surface Characterisation Using A Resonator Based Profilometer

Salam M. Harb,

Coventry University, Coventry, CV1 5FB, UK

Stuart T. Smith,

Precision Engineering Laboratory, UNC Charlotte, Charlotte, NC 28223

Abstract

The design of a simple resonator based 2-dimensional surface profilometer is presented. It is envisaged that this instrument can be employed to provide information concerning various surface characteristics such as height variations, edge point detection and surface property variations such as the ratio of elastic modulus to density. In this design, the resonator probe is used to detect the interfacial force between the probe tip and specimen with the output from this sensor providing a feedback error signal into a digital piezoelectric translator (DPT). The specimen is mounted on a carriage which traverses beneath the probe using a high precision single degree of freedom polymer bearing slideway. Motion of the slideway is monitored by a laser interferometer. Variations of the contact forces between the probe and specimen alter its resonant characteristics and this is measured using standard phase locking techniques. Under closed loop control, the output from the phase locked loop is fed back to the DPT which adjusts the probe height to maintain a constant interaction between probe tip and surface. Specimens ranging from thin metallic films to flat glass surfaces are used to illustrate its metrological capability.

1. Introduction

Resonator probes have been used for several decades in a broad range of applications which required high resolution sensors. These includes sensors for liquid, or gas, density and viscosity, liquid level, mass, and microbalance [1,2]. They have also been popular for the sensing element in scanning probe microscopy [3]. The principle of operation of these sensors is based on the fact that when a mechanical element oscillated at one of its resonant frequencies comes into near-contact or contact with another object it alters the stiffness, mass and damping of the resonator consequently causing changes in the resonance frequency. This can be detected by monitoring the phase, the amplitude or the impedance of the signal. In the experiments presented in this paper, the output signal is the demand to the voltage controlled oscillator of the PLL operating at constant phase.

This paper presents the design and results of a simple resonator profilometer using a probe that was initially designed as an analogue touch sensor for coordinate measuring machines [4]. In this design the resonator probe consist of a rod of square cross section with two PZT plates attached on opposing sides. One PZT is used as the drive and the other serves to pick-up the subsequent strain signal. The signal is then fed into a phase locked loop (PLL) analogue circuitry which drives and detects the shift in resonance frequency. In this application, the probe is used merely as a nulling device in closed loop system to profile a range of different substrates selected for their different material properties. The signal of the PLL, after passing through a PID

controller, is appropriately amplified and used as a closed loop error signal to feed a digital piezoelectric translator (DPT), which in turn is used to control the vertical displacement of the probe. A 10 μm radius diamond stylus was used. The specimen is traversed underneath the probe using a high precision single degree of freedom polymer bearing slideway. As specimen is scanned, the probe 'feels' the contact force, and moves up and down in order to maintain constant phase of the resonator. The vertical displacement is used to provide height, edge point detection and property variations of the surface.

Comparison with profiles over similar regions obtained using a Rank Taylor Hobson Talystep revealed noticeable discrepancies between the two techniques, particularly when different materials are encountered during a scan. Based on these observations, further investigations into the contact mechanism between the probe and specimen surface have been undertaken, Vidic *et al.* [6].

2. Probe design and Theory

The experimental investigation of this study was conducted using resonator sensor consists of a cantilever beam that is vibrating longitudinally. For such system the resonance frequencies of an unloaded (non contact) probe is given by [5]:

$$\omega_i = \frac{(2i-1)\pi}{2l} \sqrt{E/\rho} \quad (i=1, 2, 3, \dots) \quad (1)$$

where, E , ρ , and l are Young's modulus, density and probe length, respectively

In general any changes in forces at the end of the probe will affect the dynamic characteristics of the system and hence its resonance frequencies. The interaction between oscillating probe and contacting surfaces can be modelled as a simple longitudinal oscillator of length l , with boundary conditions of added mass m_e , stiffness, k_e , and damping, h_e at the free end. Applying Rayleigh's approximate method, it is possible to derive a simple second order model with a natural frequency given by :

$$\omega_n \approx \sqrt{\frac{k_c + k_e}{M + m_e / 3}} \quad (2)$$

where k_e and m_e are given by [6]

$$k_e = (6E^2 P_o R)^{1/3} \quad (3)$$

$$m_e \approx \frac{\pi^2 P_o R}{2E} (1 - \nu^2) \rho \quad (4)$$

and P_o is the applied normal load.

3. Profilometer Instrument

The profilometer probe consists of an aluminum rod of square cross section (5x5x10 mm) oscillated longitudinally using a piezoelectric actuator and pick-up sensor cemented on the opposite sides of the beam. The probe itself consists of a standard 10 μm diamond tip stylus screwed into the free end of the rectangular rod.

Although, the resonance frequency of the device strongly depends on its mounting arrangement, the frequency of this sensor was typically in the region of 18.5 kHz with behavior characteristic of a second order system. An analogue phase lock loop (PLL) NE 565 integrated circuit is used to oscillate the probe at one of its resonance frequencies and monitor the frequency

shift as the result of interfacial force variation. Figure 1 shows the block diagram of the control system of the resonator probe. During these experiments, the phase shift between the drive and the strain signal was set near -90° . As contact (or near contact) takes place, changes in resonance

Resonator probe block diagram

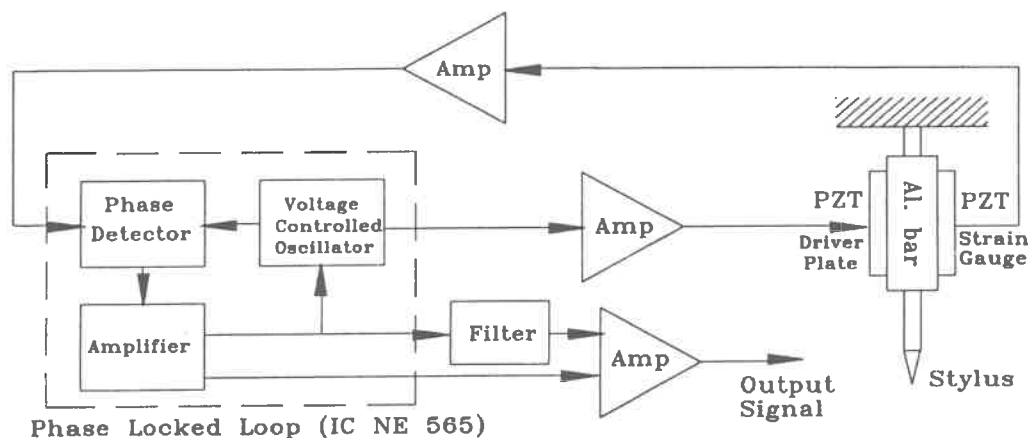


Figure 1. Schematic diagram indicating elements of the resonator probe

frequency result in a phase angle shift from its preset value. The PLL techniques adjusts the driving signal frequency to null the change in phase angle and producing a proportional output signal with a conversion rate of 0.73 mV Hz^{-1} .

For this application, the complete resonator has been mounted to the z translator, the underside of a vertical piezoelectric translator (Queensgate Instruments model S300 digital piezoelectric translator, DPT), the holder of which forms part of a z-axis coarse positioner. DPT consists of tubular piezoelectric actuator (PZT) with a capacitance sensor along the central axis to monitor the actual extension. The DPT translator has a linear extension range of $15 \mu\text{m}$ with noise level of 0.6 nm rms. at bandwidth of 1100 Hz. The z axis is mounted into a rigid mechanical bridge structure above a horizontal slideway.

A schematic diagram of the profilometer is shown in Figure 2. The changes in frequency are detected by the PLL producing an output voltage which, after being filtered and amplified, and then fed through a proportional -integral -differential (PID) controller to the z-axis servo controller. The objective of the feed-back loop is to maintain the interfacial characteristics between the stylus tip and specimen at a fairly constant level (the set static force).

The specimen was scanned beneath the probe using a high-precision polymer bearing slideway sliding over a glass ceramic (Zerodur) optically flat surface. The carriage is traversed by a DC motor through a transmission system consisting of micrometer and a hysteretic coupling. Horizontal displacement of the carriage was monitored using a Michelson based laser interferometer system.

To obtain a trace of the surface, the specimen is first secured onto the aluminum carriage, and then aligned parallel with the slideway surface using angular adjustment mechanism attached to the carriage. Probe is then brought into 'contact' with the surface until the DPT actuator is

controlled at its mid-range position under closed loop control. As carriage traverses, at approximately 1 mm s^{-1} , the PID error signals to the DPT to correct for any height/ properties variations, are recorded digitally via an ADC data acquisition board and stored in a computer for further analysis. Simultaneously, the laser interferometer system was recorded via standard RS232 interface. Recording and analyzing data was facilitated using LabView¹ and MicroSoft Excel software packages.

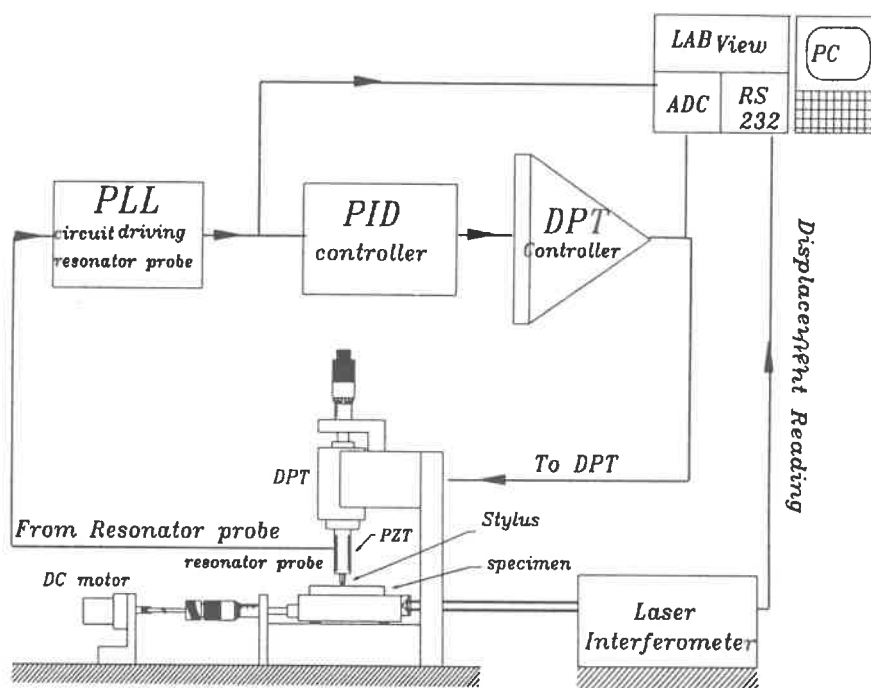


Figure 2. Schematic diagram of the resonator profilometer

4. Performance and Results

Traces over reasonably flat engineering surfaces ranging from ground and lapped steel to flat glass and silicon surfaces were investigated during these experiments. Also, to observe the influence of different material properties within a profile, some partially coated with aluminum and gold, were also measured.

Figure 3 shows a typical resonator profile over a scratched glass microscope slide surface. In this, more than 8 traces in a single traverse direction were recorded over the same line and the results were fairly consistent and reproducible. Similar test was conducted over a 100 nm step height etched on a silicon wafer, figure 4. Repeating the measurement shows similar features and, although near to the limits of resolution of this system, the step height measured corresponds closely to the expected value.

¹ National Instrument

Figure 5 shows a typical resonator profiles over a microscopic slide partially coated with thin film of aluminum, nominally 50 nm. In these traces, measurement starts when probe is in contact with glass surface and traversed to cross over the aluminum film and then traversed back to the starting point. Readings covering glass surface, aluminum film and at the transition region were recorded. As shown in figure 5 repeating the same traverse produced very similar profiles which indicates the consistency of the measurement. However, the interesting point that can be inferred from these traces is the behaviour of the probe as it encounters different surfaces. At the transition region the probe exhibits rapid changes which do not correspond to the actual height variations but more likely it is due to sudden changes in the physical properties which forces the controller to exaggerate the probe's height displacements about 4.5 μm , in order to compensate for these changes. After probe crosses over completely, the controller settles down and the variations in profile can again be attributed mainly to the surface topography. Comparing this with similar traces obtained using stylus type surface measuring instrument (Talystep), figure 6 shows that the transition between the surfaces is much smoother. Nonetheless, the magnitude of the film measured by the both instruments was not far out, taking into account that neither Talystep nor the resonator profilometer were reliably calibrated prior to these tests.

Similar traces were repeated over microscope slide coated with nominally 100 nm film of gold figure 7. Again similar results were observed indicating the transition region between the surfaces. However, the magnitudes of jumps were less evident about 0.5 μm compared to previous one (figure 5).

5. Conclusion

A prototype resonator profilometer has been designed and tested. Resonator traces over various reasonably flat engineering surfaces were recorded. and results indicate probe's response to the height variations (surface topography) and surface properties variations, of particular interest are the variation in Young's modulus and density. These were evident when surfaces having changes in materials properties were scanned. The instrument shows high sensitivity in detecting the edge point (i.e. the transition between two surfaces).

Further work is currently, underway to extend the design into 3-dimensional surface characterisation and to be able to correlate generated profiles to some useful surface properties. Such a measuring system may require incorporation of this instrument with a conventional stylus measuring device in order to differentiate topography from the physical properties discriminated by the resonator probe.

Acknowledgment

This material is based upon work supported by Coventry University Central Funding Scheme. Dr. Harb would like to thank Royal Academy for engineering and Nuffield foundation for supporting his visit to the Precision Engineering Laboratory at The University of North Carolina at Charlotte. As well, Mr. G. Singh, and R. Edgeworth of Precision Engineering group at UNCC Charlotte, for their help and support.

References

- [1] Langdon R.M. Resonator sensors-a review, *Review article, J. Phys. E: Sci. Instrum.*, **18**, 1985, 103-115.

- [2] Morten, B., De Cicco, G., Prudenenziati, M. Resonant pressure sensor based on piezoelectric properties of ferroelectric thick films, *Sensors and Actuators A*, **31** (1992), 153-158.
- [3] Wienmann, M. Raduis, R. and Assmus, F. Measuring profile and position by means of vibrating quartz resonators used as a tactile and nontactile sensors, *Sensors and actuators A*, **37-38** (1993) 715-722.
- [4] Harb, S.M., Vidic, M. Resonator-based touch-sensitive probe, *Sensor and Actuators*, **A50** (1995), 23-29.
- [5] Smith, S.T. Chetwynd D.G. *Foundations of Ultraprecision Mechanism Design*, Gordon and Breach, (1992), NY.
- [6] Vidic, M., Harb, S.M., and Smith, S.T., Observation of contact measurements using a resonance based touch sensor, *Precision Engineering*, submitted for publication

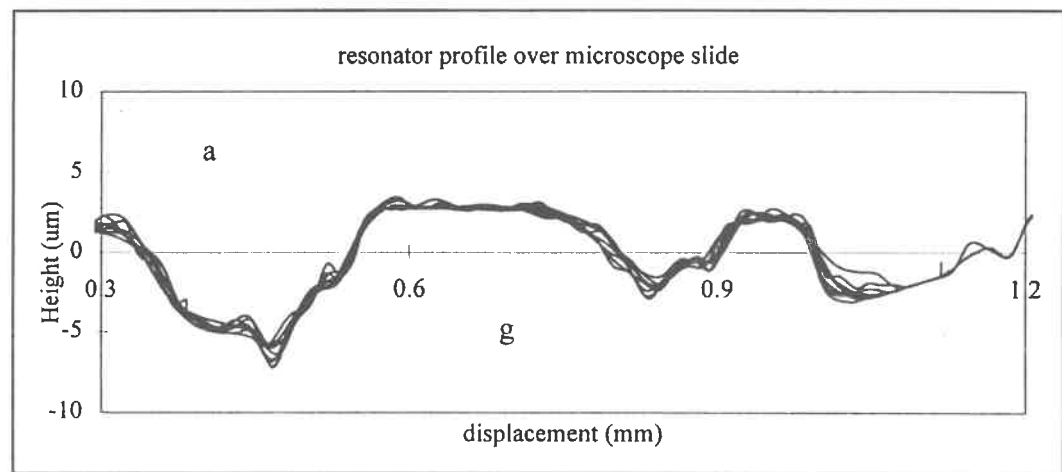


Figure 3 Typical traces over scratched microscope slide glass surface

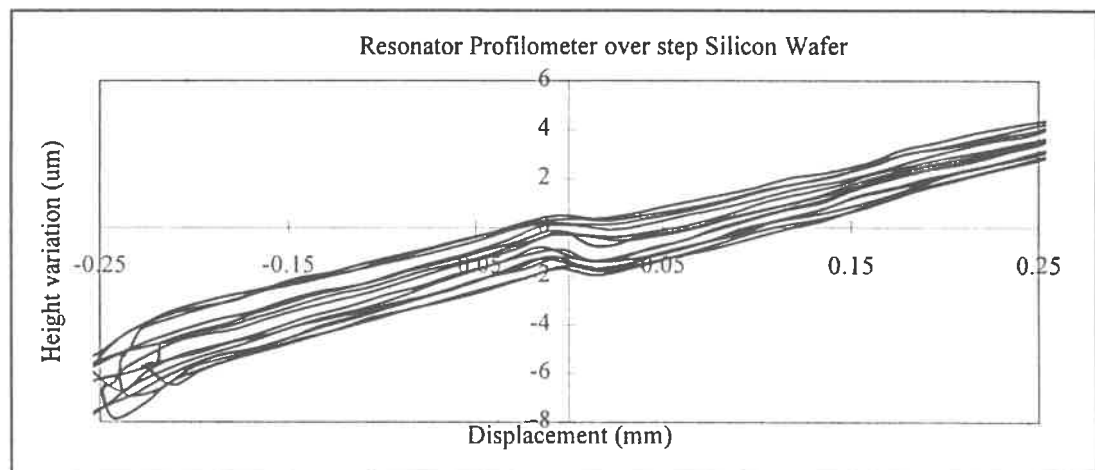


Figure 4 Typical resonator profile over a step height on a silicon wafer

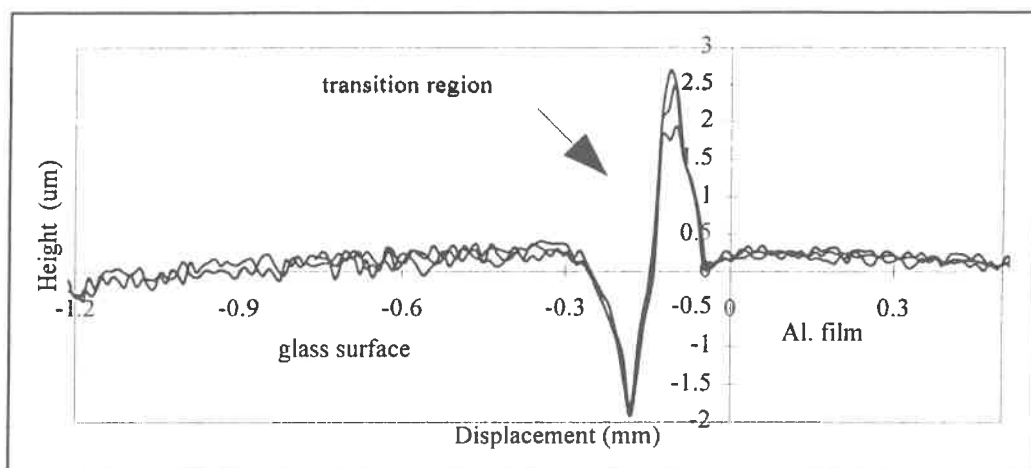


Figure 5 Profiles over a microscopic slide partially coated with 50 nm film of aluminum.

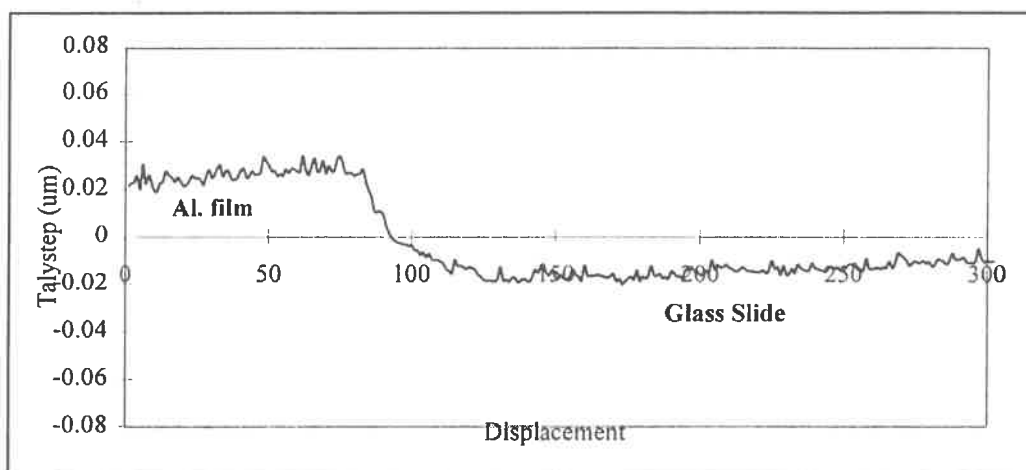


Figure 6. TalyStep trace over the same specimen in figure 5.

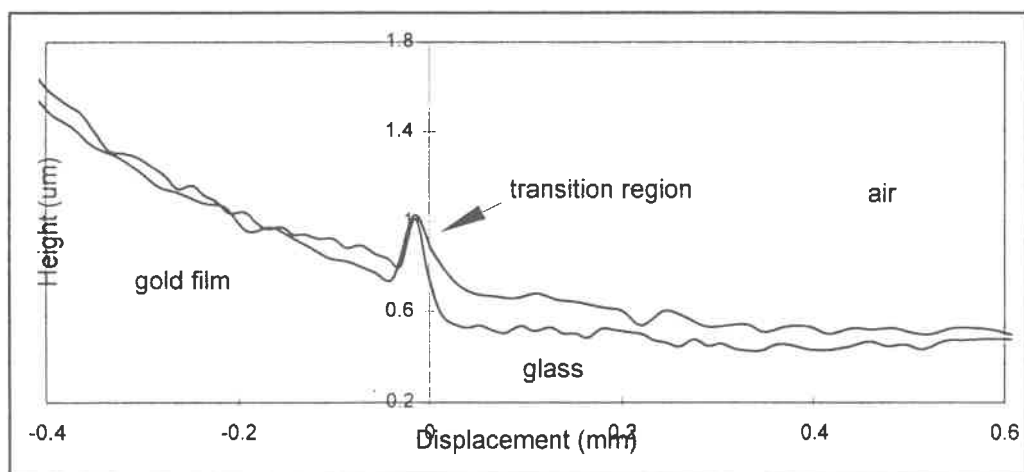


Figure 7. Resonator profile over 100 nm film of gold over glass surface

FILTERING OF SURFACE PROFILES, PAST, PRESENT AND FUTURE

Xiangyang Liu and Jay Raja

Center for Precision Metrology
University of North Carolina at Charlotte
Charlotte, NC 28223

ABSTRACT

The evolution of various filtering techniques used in surface metrology is presented in this paper. The characteristics of these filters and a brief comparison is also given. Wavelet filter banks are investigated as potential filtering technique for surface texture analysis.

INTRODUCTION

Filtering is widely used in surface texture characterization for separating the different wavelength components in a surface profile. The earliest filter used in surface measurement is the analogue RC filter^[1]. Advances in signal processing and the availability of inexpensive computing have led to the use of digital filters in recent years. Compared with analogue filters, digital filters have many advantages and are preferred in many applications because of the flexible characteristics, stable performance, wide frequency range and high accuracy^[2].

FILTERS - PAST

The earliest filters used in surface texture are based on a two stage capacitance-resistance network^[3]. The transmission characteristic of 2RC low pass filter is given in Figure 1(a), λ_c is the long wavelength roughness cutoff and at this wavelength the filter provides a 75% transmission. The introduction of microprocessor led to the implementation of this filter in a digital form. The implementation for 2RC filter involves either a convolution method or a recursive method. The main disadvantage of a 2RC filter is the nonlinear phase which causes distortion of the filtered signal. In addition the 75% transmission at the cut off is not convenient for analysis of waviness. If the transmission at the cutoff is 50% the sum of roughness and waviness signal can provide the unfiltered signal. The phase distortion of the 2RC filter could be eliminated with a symmetrical weighting function. However the need for 50% transmission and phase correct filter lead to the Gaussian filter with a 50% transmission curve which is the filter currently used.

CURRENT FILTERS

To produce an undistorted filtered profile, the constituent sinusoidal components making up the profile signal should be not shifted relative to each other during the filtering process. Phase-correct filter is a filter which has the linear phase characteristics. This can be achieved by a filter which has an impulse response symmetric about an axis of $x=x_0$. Various linear phase filter have been investigated comprehensively by Whitehouse^[4] and Raja^[5].

The weighting function of Gaussian filter is a Gaussian function^[3]. The mean line is determined for any point of the measured profile by taking a Gaussian weighting function average of the adjacent points. The cutoff λ_c is the wavelength of a sinusoidal profile for which 50% of the amplitude by the profile filter. The transmission characteristic for low-pass Gaussian filter is shown in figure 1(b). Gaussian filter has a linear phase response and there is no distortion in the filtered profile. This filter can be implemented by a convolution method or a FFT method. A fast and reliable convolution algorithm for Gaussian filter has been derived by Krystek^[6].

Profiles with deep valleys tend to pull the mean line to a lower line because of the depth of the valleys. Plateau honing produces a surface that generates such valleys. In order to overcome the effect of deep valleys a filtering procedure was developed and is currently used for calculating the Rk family of parameters. The details of the procedure can be found in DIN 4776^[7].

The national and international standards now specify Gaussian filter with a 50% transmission at the cutoff as a standard. In addition to this the standards now also specify a short wavelength cutoff λ_s to establish a defined bandwidth, as shown in figure 2. The ratio of λ_c/λ_s can be 300:1 or 100:1. The use of defined bandwidth have helped in comparison of parameter values from different instruments.

FILTERS - RECENT DEVELOPMENTS

The digital implementation of filters generate undesirable end effects if excessive form error is present. A Spline filter is suggested by Krystek^[8] for primary profile and is envisaged to cover roundness and other form characteristics. The Spline filter has the advantage over the conventional phase filter, that the edge of the measured profile are still usable. This is important especially in the case of form filtering.

Motif filter^[9] is another kind of filter used in surface metrology especially in France. In Motif method, only selected parts of the profile like the extreme peaks or valleys contribute to the extraction of roughness and waviness.

As shown in figure 2, the surface signal can be separated into form error, waviness and roughness by band pass filtering. In signal processing this type of filtering is referred to as filter banks^[10]. The use of filter banks permits bands more than the three (roughness, waviness and form error) stated above. It may be beneficial to use non uniform filter banks to relate the profile in each band to specify process variables. As an octave filter banks, wavelets can be used for this purpose, as illustrated in figure 3^[11]. The fundamental idea behind wavelet analysis is to analyze signal according to scale, i.e. process signal at different scale or resolution^[12]. Wavelet analysis provides a flexible time-frequency window localized on time-frequency plane. Therefore we have a set of windows that change in proportion to the scale of the local event observed and whose frequency bandwidth also match the frequency range of the phenomena studied. In view of signal processing, the dyadically dilated wavelets constitute a bank of octave band pass filters and the dilated scaling functions form a lowpass filter bank. The operation of discrete wavelet transform

can thus be interpreted as filtering the signal with a bank of band pass filters, followed by sampling at the respective Nyquist frequencies. By adding higher frequency bands, one adds details, or resolution, to the signal. Hence we can also get the output of a lowpass filter bank by successive approximation of the signal. It's distinct in that the wavelet bases have time-widths adapted to their frequency. We can start with looking at relatively rough structures of a signal at a low resolution and then add in finer and finer structures into our interpretation by increasing the resolution correspondingly. This method appears to have great potential for analyzing the multiscale features in engineering surfaces due to the properties of good time-frequency localization and flexible time-frequency resolution of the discrete wavelet transform .

The primary study about the application of wavelet filter banks for analyzing multiscale engineering surfaces are presented by Chen, Liu and Raja^[13,14]. In order to analyze engineering surface texture, multiscale approximation of original texture at different resolution level will be checked first. This will give us a clear overview of multiscale features of surface texture. Then the separation of multiscale surface features based on wavelength, i.e., form error, waviness and roughness or manufacturing process are made according to the information provided by the multiscale approximation. For special specimens, interested features at reasonable resolution level can be analyzed and characterized. Figure 4 illustrates the analysis of a turned surface by Daub8. From the multiapproximation, a cutoff length of 1.820 mm (corresponding to level 4) is good for separating the longer wave from the shorter wave and a 0.114 mm cutoff (level 8) separates the periodical components from the irregularities reasonably well. The approximations up to level 4 give the general trend of the profile which is caused by the variation of the machine condition in the manufacturing process. Level 5 to level 7 of the approximations, the characteristic tool mark formed by clean turning tip is well reconstructed. Above level 7, the irregular structures arising from the degeneration of the tool are gradually obtained. There are many wavelet families for various applications^[15]. The selection of appropriate wavelet is dependent on the signal analyzed.

REFERENCE

1. D. J. Whitehouse and R. E. Reason, 'The equation of the mean line of surface texture found by an electric wave filter', Rank Organization, 1965.
2. E. C. Ifeachor and B. W. Jervis, *Digital Signal Processing, a practical approach*, Addison Wesley, 1996
3. ASME B46.1-1995, Surface texture.
4. D. J. Whitehouse, 'Improved type of wavefilter for use in surface-finish measurement', *Proc Instn Mech Engrs*, 1967-68, 3K.
5. J. Raja and V. Radhakrishnan, 'Digital filtering of surface profiles', *Wear*, Vol. 57, 1979.
6. M. Krystek, 'A fast Gaussian filtering algorithm for roughness measurements', *Precision Engineering*, Vol. 19, No. 2/3, 1996.
7. DIN 4776, Parameters R_k , R_{pk} , R_{vk} , M_{r1} , M_{r2} for describing the material portion in the roughness profile.
8. GPS, Surface texture: Metrological characterization of phase correct filters-Spline filters, *Draft*.
9. ISO 12085, GPS, Surface texture: Profile method-Motif parameters.
10. G. Strang and T. Nguyen, *Wavelets and Filter banks*, Wellesley Cambridge, 1996

11. N. J. Fliege, *Multirate Digital Signal Processing, multirate system, filter banks, wavelets*, John Wiley & Sons, 1996
12. S. G. Mallat, 'A theory for multiresolution signal decomposition: the wavelet representation', *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 11, No. 7, July, 1989.
13. X. M. Chen and J. Raja, S. Simanapalli, 'Multi-scale of engineering surface', *Int. J. Mach. Tools Manufacts*. Vol. 35, No. 2.
14. X. Y. Liu and J. Raja, 'Analyzing engineering surface texture using wavelet filter', *SPIE* Vol. 2825
15. I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA, 1992. Notes from the 1990 CBMS-NSF Conference on Wavelets and Applications at Lowell, MA.

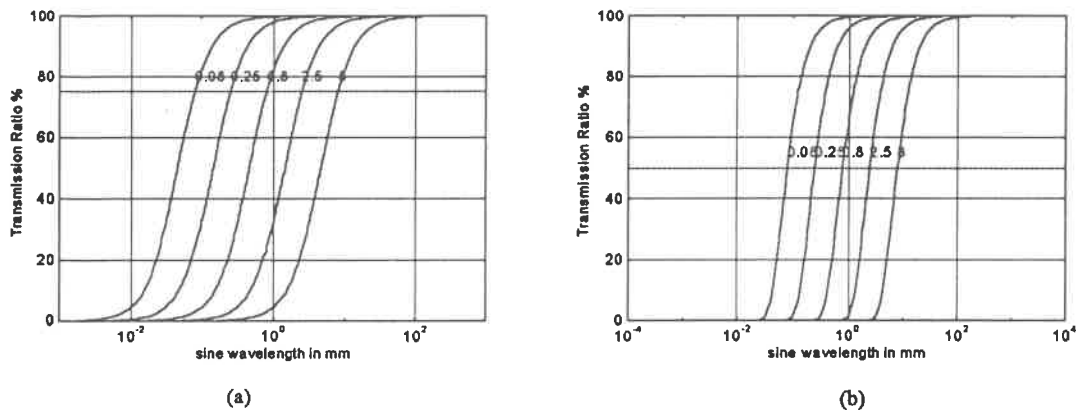


Figure 1 Transmission characteristic of (a) 2RC filter; (b) Gaussian Filter

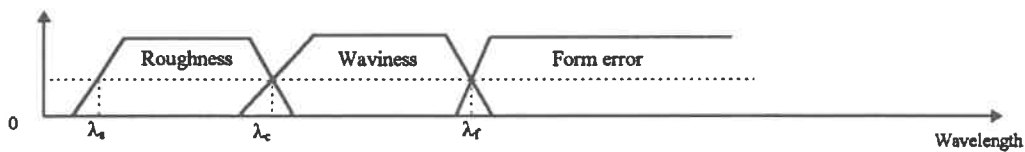


Figure 2 Roughness, waviness and form error

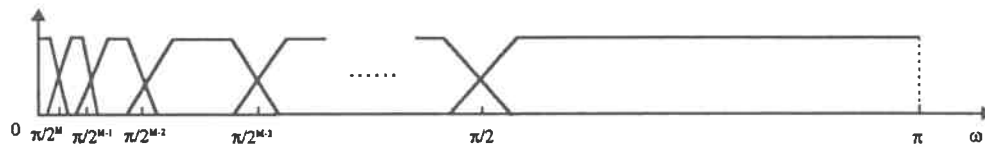


Figure 3 Wavelet filter banks

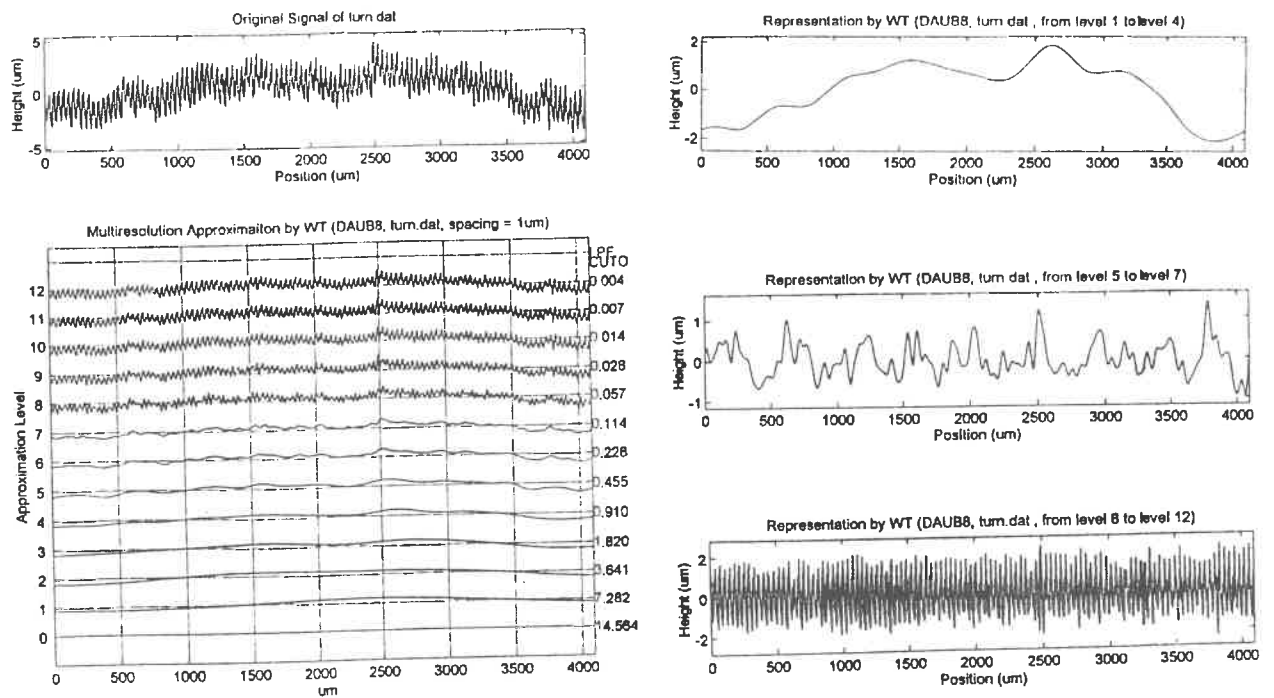


Figure 4 Analysis of the turned surface by wavelet filter

Estimation of arithmetic mean wavelength for the areal topographic data

Kazuhisa YANAGI, Yoshikazu KOBAYASHI and Seichiro HARA

Nagaoka University of Technology

Nagaoka-city, 940-21 JAPAN

1. Introduction

Areal topographic data of an engineering surface are sampled usually at regular grid points in the horizontal plane by either mechanical contact method or noncontact one. A set of the areal data sometimes includes the high-frequency irregular components. This leads to difficulty in identifying the significant summits or pits on the original data. In such a case, local slope of a facet is also ambiguous. And a certain filter is used to eliminate those noise-like components. Nevertheless it seems to be difficult to assess the mean wavelength for the areal topographic data. The 3-D spatial parameters of surface texture such as undulation per unit length/area and mean wavelength have scarcely been appeared in documentary records. It is really an elusive problem to define the geometrical wavelength or lateral asperity spacing for the areal topographic data.

Once the asperities are detected, the areal asperity density is automatically obtained. However, the arithmetic mean wavelength can not be determined because the asperities are not always regularly scattered over the surface. It is an objective of this study to effectively estimate the mean surface wavelength for 3-D surface topographic data by knowing the locations of summits and pit bottoms extracted.

2. Conventional parameters

Lateral mean spacing of 2-D profile irregularities or arithmetic mean spacing of the zero-crossings is defined as an international roughness standard (ISO DIS 4287-1,2), which is denoted by RS_m or PS_m . The mean spacing of profile irregularities corresponds to the averaged length of cycle of profile peaks and valleys. It may be recognized, however, that the mean spacing is easily affected by the high-pass filter adapted and calculation procedure. Threshold band is sometimes used to eliminate the insignificant peaks and valleys.

On the other hand, the mean surface wavelength in two-dimensions has been proposed by regarding a profile as an ideal sinusoidal wave form of constant amplitude. It is given by the following equation.

$$\lambda_a = 2\pi R_a / \Delta_a \quad (1)$$

where, R_a is the averaged profile height. Δ_a is the averaged profile slope.

It may be possible to replace R_a and Δ_a with R_q (root-mean-square profile height) and Δ_q (rms profile slope). In that case, another notation and parameter modification will be required.

By expanding such an idea to the areal surface topographic data, the similar formula can be introduced. For a sinusoidal surface form, the 3-D mean wavelength can be expressed by eq.(2).

$$S\lambda_a = 2\pi SR_a / S\Delta_a \quad (2)$$

where, SR_a is the averaged deviation of amplitude and $S\Delta_a$ is the averaged value of absolute local slopes.

SR_a is rather a stable parameter. However, SA_a is changeable and influenced by the calculation procedure. In this study, the areal smoothing and differentiation technique⁽¹⁾ was used for the slope calculation at every sampling point.

3. New spacing parameter based on the location of summits and pits

We tried to propose a new definition of the lateral mean wavelength for the areal topographic data by revealing the locations of effective summits and pit bottoms. First of all, the mean plane derived from a regression curved surface is set to be a base plane by which the convexity and concavity can be judged. The asperity summits below the base plane and the pit bottoms above the base plane were not taken into account in this study. Those significant summits and pit bottoms can be extracted by following the smoothing and differentiation techniques. A square patch area $W_x \times W_y$ by which the summit or pit can be identified was introduced⁽²⁾ and scanned over the topographic data as shown in Fig.1. This means that at most one summit lie in the middle of the area and the minimum distance between two adjacent summits is half a side of the square.

It is expected that a zero-crossing point exists between such a summit and a pit bottom as shown in Fig.2. However, there are a number of combinations of the adjacent summit and pit bottom, and so we pick up only the nearest combination for each summit or pit and substitute it for the 3-D surface wavelength as shown in Fig.3. We thought the arithmetic mean value of the distance between a summit and its nearest pit bottom can be correlated with the lateral mean wavelength of 3-D surface topographic data. We propose an assessment parameter given by the following equation for the mean wavelength of the areal surface topographic data.

$$DS_m = \kappa \overline{d_{scmin}} \quad (3)$$

where, $\overline{d_{scmin}}$ is the arithmetic mean distance for all the combinations of nearest summit and pit bottom, and κ is a constant ($\kappa=2.0$ in this study).

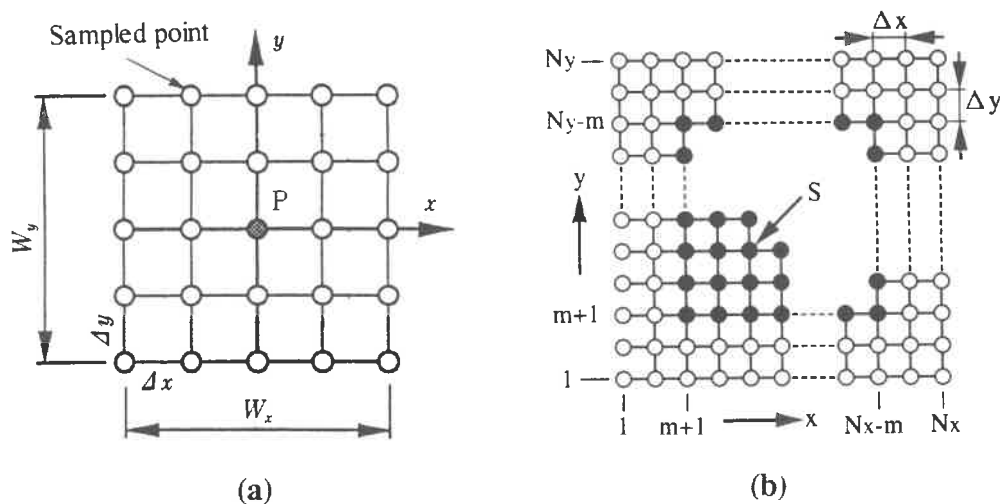


Figure 1 Square patch area to identify a summit or a pit bottom

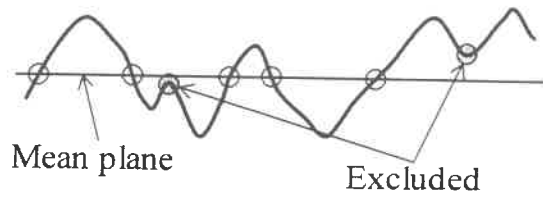


Figure 2 Zero-crossing between a summit and its neighboring pit bottom

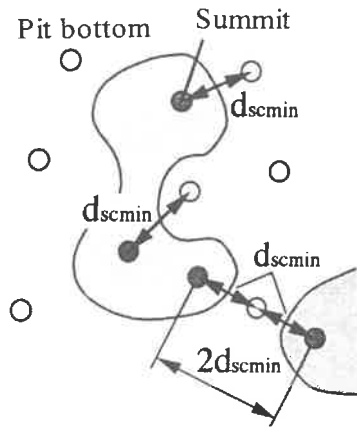


Figure 3 Indirect measures for 3-D surface wavelength

4. Discussions

Let us suppose the extracted summit number be N_s and the measurement areal be A . Then the areal density of the summit becomes $\eta = N_s/A$. When the summits above the mean plane are concerned, the density is expressed by η_p . If the summits are distributed regularly in the lattice form, the lateral mean spacing among the summits can be obtained from a geometrical analysis. Figure 4 shows the square lattice distribution of the summits and equilateral triangle one. If the distribution is in such a manner as k -angle regular polygon, the mean spacing between adjacent summits can be expressed by the following equation.

$$D_a = \frac{k}{\pi} \int_0^{\frac{\pi}{2}} \frac{D_0}{\cos \theta} d\theta = \frac{kD_0}{\pi} \left[\ln \left| \tan \left(\frac{\pi}{2k} + \frac{\pi}{4} \right) \right| \right] \quad (4)$$

where, D_0 is reference width for the area in which a summit has a share.

In case of the square distribution, $D_a(\text{Square}) = 1.122D_0$ and $D_0 = 1/\sqrt{\eta_p}$. And for the equilateral triangle, $D_a(\text{Triangle}) = 1.049D_0$ and $D_0 = \sqrt{3}/\sqrt{2\eta_p}$. It should be noted here again that only the summits above the mean plane are taken into consideration.

It is an interesting work to compare D_{sm} with D_a for a variety of machined surfaces. In most cases, D_{sm} becomes possibly larger than D_a . When the summit density is relatively small and the summits are distributed separately, D_{sm} might be smaller than D_a . This means that the macroscopic convexity is fundamentally composed of a few summits and the distance between two adjacent convexities is longer. The experimental results are tabulated in Table 1. It

can be seen that D_{Sm} is larger than $D_a(\text{Triangle})$ for all the samples. And so, D_{Sm} is given by eq.(3) in this paper seems to be reasonable parameter for the mean wavelength. Contrary to this, $S\lambda_w$ values are much larger than D_{Sm} for all the samples. This is attributed to the different physical meanings of D_{Sm} and $S\lambda_w$. Since $S\lambda_w$ is inversely proportional to the averaged local slope, it is greatly influenced by the calculation procedure of the slope. On the other hand, D_{Sm} depends on the definition of summit or pit bottom.. Further discussions on the local slope as well as the asperity summit will be necessitated to assess the mean surface wavelength more property.

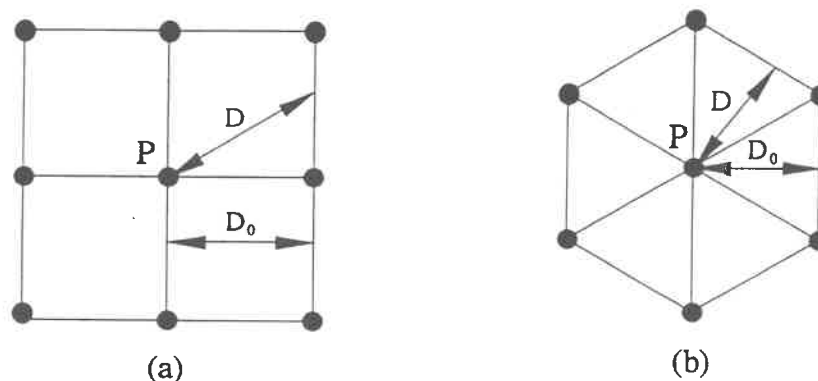


Figure 4 Equilateral distributions of the effective summits

Table 1 Experimentally obtained parameter values for the samples under test

Samples		η_p μm^{-2}	$D_a(\text{square})$ $\mu\text{m}, \text{nm}$	$D_a(\text{Triangle})$ $\mu\text{m}, \text{nm}$	D_{Sm} $\mu\text{m}, \text{nm}$	$S\lambda_w$ $\mu\text{m}, \text{nm}$
Lapping film	LF1	0.118	3.26 μm	2.84 μm	3.81 μm	7.20 μm
	LF2	0.092	3.69	3.21	4.27	7.78
	LF3	0.096	3.62	3.15	4.78	8.73
EDM	ED1	0.019	8.12	7.06	9.14	15.6
	ED2	0.009	11.9	10.4	17.0	25.2
Base film (optical)	BF1-O	0.372	1.84	1.60	1.79	2.86
	BF2-O	0.150	2.89	2.51	3.49	5.15
	BF3-O	0.328	1.96	1.70	1.87	2.86
Base film (SEM)	BF1-S	0.204	2.48	2.16	3.21	4.78
	BF2-S	0.135	3.05	2.65	4.32	6.17
	BF3-S	0.191	2.56	2.23	3.08	4.74
Head slider	HS1	22.0	239 nm	208 nm	282 nm	510 nm
	HS2	18.8	258	225	274	308

5. Concluding remarks

We tried to estimate arithmetic mean wavelength for the areal topographic data by knowing the spatial distributions of significant summits and pit bottoms. All the nearest combinations of summits above the mean plane and pit bottoms below the plane were taken into consideration. The arithmetic mean value of those nearest distances was regarded as a measure of the mean surface wavelength. To popularize the idea of the 3-D mean wavelength, it is essential to disclose the discrepancies among those related parameters and to develop more practical computer programs for calculation. The authors expect a certain wavelength parameter to be included in the 3-D roughness parameters of ISO standards in the near future.

References

- (1) K.Yanagi and N.Kobayashi : An assessment of tribological surface function through quantitative characterization of surface topography, Trans. Mat. Res. Soc. Jpn.,(1994) 923-928.
- (2) K.Yanagi and N.Kobayashi : Surface Topography Assessment for Information Storage Media, Proc. 3rd Int. Conf. On Ultraprecision in Manufacturing Eng.. Aachen, (May 1994) 382-385.
- (3) P.J.Clark and F.C.Evans: Distance to nearest Neighbor as a Measure of Spatial Relationships in Populations, Ecology, 35, 4,(1954)445-453.

EVALUATION OF FUNCTIONAL FEATURES OF ORTHOPAEDIC JOINT PROSTHESES SURFACES USING WAVELETS

X. Q. JIANG L. BLUNT K. J. STOUT

*Centre for Metrology
School of Manufacturing and Mechanical Engineering
The University of Birmingham
Birmingham B15 2TT, UK*

1. INTRODUCTION

The surface topography of the counterface, femoral head rubbing against the UHMWPE acetabular cup has been recognised as one of the most important factors affecting the functional performance of the joint replacement system. A great deal of research indicates that wear rate of the polymer is determined primarily by the surface topographic characteristics of the femoral head [1-7]. Hall [1] has shown that statistical roughness parameters of the retrieved femoral heads (37 explanted) are significantly greater when compared with values from newly prepared prostheses. Bauer [5] found that if the initial surface roughness of new femoral heads was smoother than standard manufacture specifications, the retrieved cobalt-chrome heads had a low surface roughness and less deep scratches than those of heads with standard roughness values. Fisher [3] showed that the wear of the UHMWPE was greatly affected by the higher counterface roughness, and he points out that it is essential to control the topography of the counterface of the femoral head to ensure low wear rates in service. Also it has been reported that an increase in the roughness of the surface of the femoral head, R_a , from $0.01\mu\text{m}$ to $0.1\mu\text{m}$ will cause a 13 fold increase in the wear rate of the femoral head [7]. However, from the functional view point, all surface features are very significant, for example, multi-scalar peaks of bearing surface directly affect the prostheses wear. During operation of the surface, these peaks will act as sites of high contact stresses and abrasion. Consequently, wear particles and debris generated by these peaks will be liberated rapidly into the surrounding tissues of the joint, causing adverse cellular reactions which lead to bone resorption and subsequent loosening of the joint. Therefore, these multi-scalar peaks will impact directly on the biomechanical and physical properties of the joint replacement system.

The purpose of this research is to evaluate the peak features of bearing surface of orthopaedic joint prostheses by using wavelet digital analysis, with an emphasis on linking the identification of the peak distributions in different scales. In this paper, a two-dimensional (2D) wavelet digital filtering technique is employed for addressing multi-scale properties of the peaks, localising position of these peaks and evaluating corresponding bearing properties of the femoral head. Furthermore, measurement of metal heads, obtained by from different manufacturing processes has been conducted to demonstrate the applicability of using wavelets in the analysis of these surfaces.

2. WAVELET ANALYSIS FOR IDENTIFICATION MULTI-SCALAR PEAKS

Generally, the 3D surface topography of orthopaedic joint prostheses is superimposed by a series of frequency components:-

$$z(x, y) = \eta(x, y) + \eta'(x, y) + \eta''(x, y) + \xi(x, y) \quad (1)$$

where, $\eta(x, y)$ stands for roughness with high frequency, $1/\lambda_r \sim 1/\lambda_r$, $1/\lambda_r$ indicates high frequency limit by sample interval, δ , and $1/\lambda_r$ is roughness frequency limit. $\eta'(x, y)$ and $\eta''(x, y)$ are "waviness" and "form error" with low frequencies (waviness $1/\lambda_r \sim 1/\lambda_w$, $1/\lambda_w$ is waviness frequency limit and form $< 1/\lambda_w$) respectively.

Whilst the $\xi(x, y)$ is composed by multi-scale peaks, pits and scratches, whose wavelength would cover a wide frequency range ($1/\lambda_s \sim 1/\lambda_w$).

In order to separate and identify the distributive features of peak, the 3D original measured signal $z(x, y)$ is decomposed into a space-scale plane. Mathematically, a 2D discrete wavelet transform $W_{j,k}(z)$ can be expressed by the inner product of 3D raw data $z(x, y)$ and a 2D bases wavelet $\psi_{j,k}(x, y)$

$$W_{j,k}(z) = \langle z(x, y), \psi_{j,k}(x, y) \rangle \quad (2)$$

where, $\psi_{j,k}(x, y)$ stands for the discrete prototype wavelet or mother wavelet which has compact support or fast decay in both the time and frequency domains and whose integral is zero. $j, k \in \mathbb{Z}$: j is a scale parameter which can be used to illustrate a dynamic transmission bandwidth, k is a translation parameter. The main approach to wavelets uses the two-channel analysis filter banks, and wavelets developed from the filter coefficients [8-9]. However due to the frequency responses of the two-channel filter, lowpass H_0 and highpass H_1 , are not ideal brick wall filters, there are overlaps that lead to aliasing, amplitude modification and phase distortion. In recent study, another two-channel synthesis filter F_0 and F_1 was specially designed to compensate the error of the analysis filter. When the synthesis bank is the inverse of the analysis bank, the biorthogonal wavelet filtering pair is formed. Therefore, the various information from the surface topography can be separated quantified accurately.

With the 2D discrete biorthogonal wavelet transform [10], the surface topography is expanded as a wavelet series, according to a transmission bandwidth ($2^{-j} = (1/\lambda_s \sim 1/\lambda_w)$).

$$W_{j,k}(z) = (Ca_{j,k}, Cd_{1,k1}, \dots, Cd_{j,k}) = \langle z(x, y), \tilde{\psi}_{j,k}(x, y) \rangle \quad (3)$$

where, $\tilde{\psi}_{j,k}(x, y)$ is the two-dimensional analysis wavelets, $Ca_{j,k}$'s are approximation coefficients that represent the output of lowpass filter bank, whilst the high frequency bank ($1/\lambda_s \sim 1/\lambda_w$) is indicated by the corresponding detail coefficients, $Cd_{1,k1}, Cd_{2,k2}, \dots, Cd_{j,k}$, which represent the frequency components of $\eta(x, y) + \eta'(x, y) + \xi(x, y)$. Since the peak points of wavelet detail coefficient along the space-scale plane record the multi-scale peak properties [10], a wavelet thresholding technique is specially designed to diagnose peaks position and amplitude, assuming that each amplitude of the multi-scale peaks in the different scale is larger than two times of the standard deviation of the corresponding detail coefficient. A hard thresholding estimator is given by:-

$$Cd_j^{hard} = \begin{cases} 0, & Cd_j < 2\sigma_j \\ Cd_j, & Cd_j \geq 2\sigma_j \end{cases}, \quad \sigma_j = \sqrt{\frac{1}{MN} \sum_{i=1}^N \sum_{r=1}^M cd_j^2} \quad (4)$$

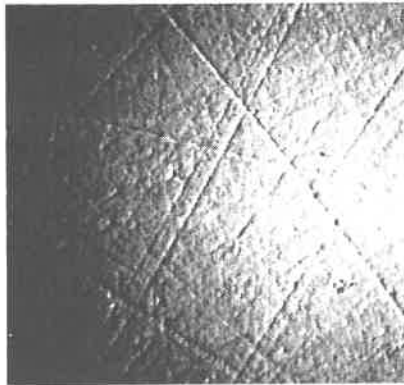
If a detail coefficient is smaller than the $2\sigma_j$, the coefficient should be replaced by a zero, whilst in the case where a detail coefficient is larger than or equal to the $2\sigma_j$ the coefficient is retained. After this processing, the detail coefficients which only represent multi-scale peak information, are obtained. The multi-scale peaks can then be reconstructed by using the 2D inverse discrete wavelet transform, with representation of details $D_j(x, y)$

$$\sum D_j(x, y) \Leftarrow \sum \sum Cd_j^{hard} \psi_{j,k}(x, y) \quad (5)$$

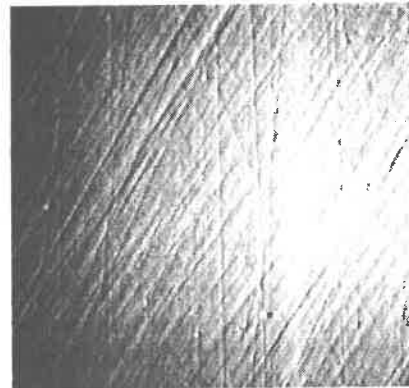
where, $\psi_{j,k}(x, y)$ is the two-dimensional synthesis wavelets.

3. EVALUATION OF FUNCTIONAL FEATURES

Fig.1 and Fig.2 show the typical bearing topographies of the two femoral heads obtained from two manufacturing processes. The Fig.1(a) and Fig.2(a) refer to the polar regions of the heads, whilst Fig.1(b) and Fig.2(b) refer to the equatorial regions of the heads. The 3D surface data from experiments was originally obtained by means of a WYKO TOPO 3D phase shifting interferometer with a sample interval of $1\mu\text{m}$. The sample Matrix was 242×233 . It can be seen clearly from the figures that the superpolished topography looks 'fine and smooth' except for a few isolated pits, peaks and some deeper scratches. The surface in Fig.1 has the smoother topography than those of Fig.2. The scale of the features is clearly greater in Fig.2, where larger peaks, pits and scratches can be seen.

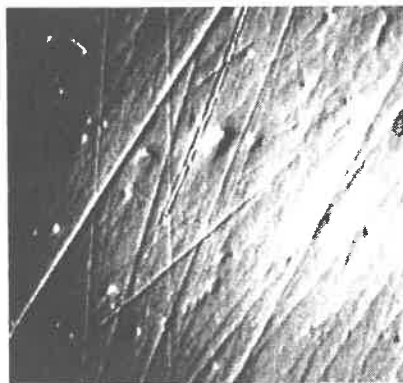


(a) in the polar region

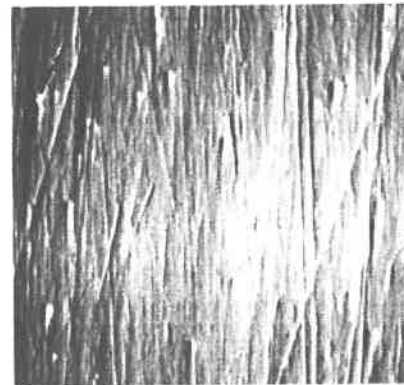


(b) in the equatorial region

Fig.1 bearing surfaces of a femoral head (1)



(a) in the polar region



(b) in the equatorial region

Fig.2 bearing surfaces of a femoral head (2)

Using a 2D discrete biorthogonal wavelet transform and the wavelet hard thresholder, identification of the peak distributions from 3D original data was carried out. Fig.3 and Fig.4 shows the results. The corresponding bearing curves of different details can be seen in Fig.5 and Fig.6.

As shown in the Fig.3 and Fig.4, it can be clearly seen that the form and waviness within the 3D prostheses bearing surface have not only been removed, but the roughness, pits and scratches have also been separated.

The multi-scale peaks are recovered perfectly and their peak positions retained accurately. These peaks belong to a frequency region of $1/0.001 \sim 1/0.032$ (1/mm).

Although there is different distribution of the multi-scale peaks shown in Fig.3 and Fig.4, the bearing curves produced by the peaks of the different details seem to have an approximately same shape which is indicated in Fig.5 and fig.6. The 35~40% bearing area is covered by the multi-scale peaks. Fig.5 show that the peak height values are less than $0.016\mu\text{m}$, whilst in Fig.6 they are $0.071\mu\text{m}$. It can be shown that the amplitude distribution of the peaks of the different details in the transmission band differs significantly. In the scale $j = 1 \sim 3$, there is a slight increase in the amplitude of bearing curves of different details as defined from the equation (5), thus the peak amplitude produced by these details does not have a remarkable influence on the initial stages of operation. Nevertheless, in the scale $j = 4 \sim 5$, bearing curves of the details change exponentially, it is quite clear that the main amplitude value of multi-scale peaks is produced by the details at scale $j = 4 \sim 5$. In this case, the corresponding bearing properties are poor, and lead to poor contact area and high contact stress within an 242×233 area. From a functional application point of view, wear particles and debris will rapidly be generated by such multi-scale peaks, during initial stages of the working cycle. Therefore, the bearing information of detail in different scales, can be used to describe contact stress, loaded area and asperity volume during initial stages of wear properties.

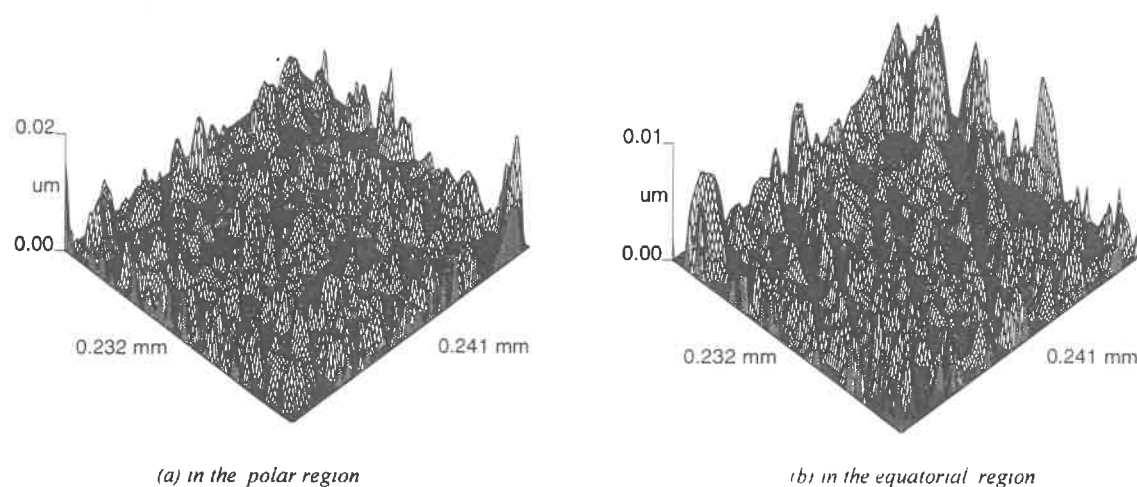


Fig.3 peak distributions of bearing surfaces in the Fig. 1

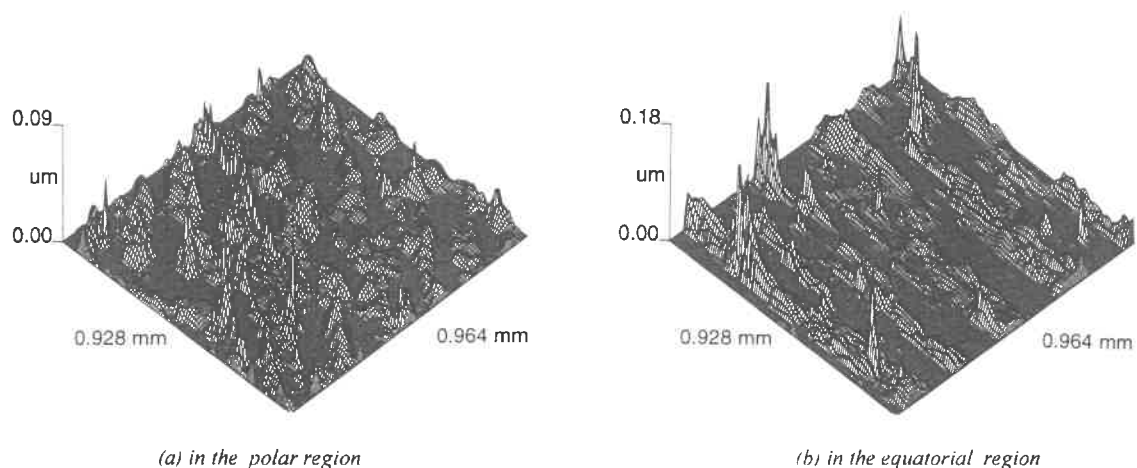
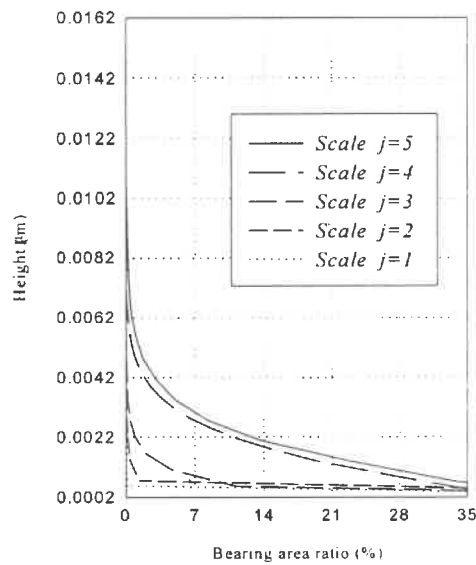
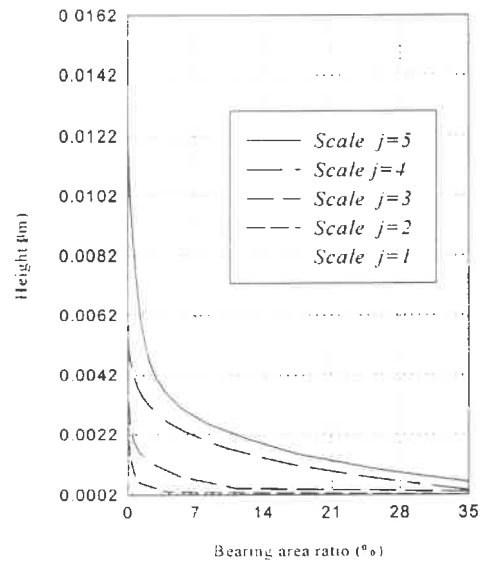


Fig.4 peak distributions of bearing surfaces in the Fig. 2

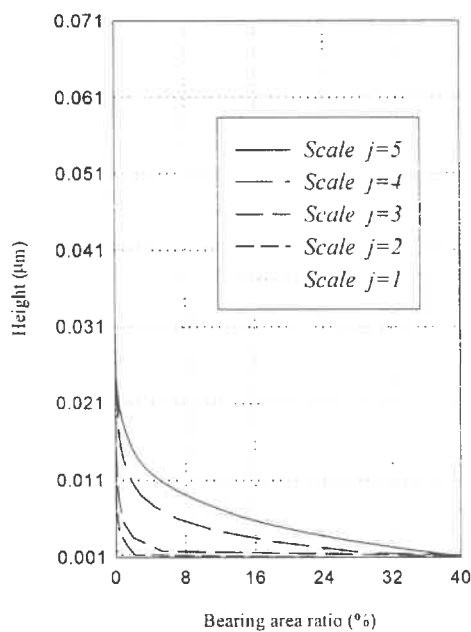


(a) in the polar region

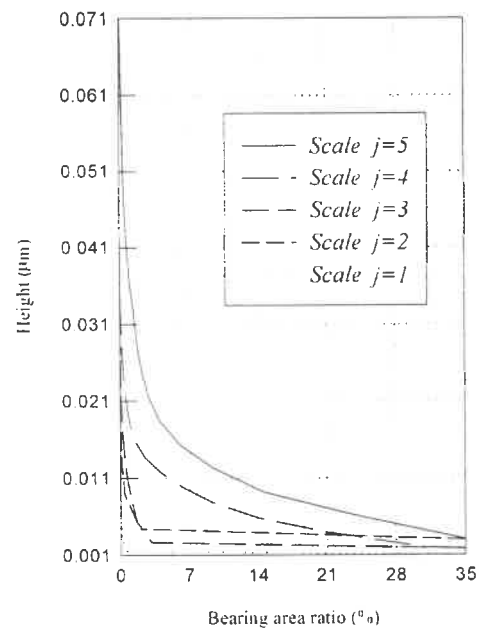


(b) in the equatorial region

Fig.5 Bearing curves of the peak distribution of the different details in Fig.3



(c) in the polar region



(d) in the equatorial region

Fig.6 Bearing curves of the peak distribution of the different details in Fig.4

4. CONCLUSIONS

By means of wavelet digital analysis techniques, the multi-scalar peaks the 3D surface topography of orthopaedic joint prostheses can be identified and extracted accurately. From the research results obtained, the bearing area ratios of different peak patterns have been analysed, and the regular pattern of peak distribution changed with the transmission band employed. In the light of such results, the friction, contact stress and loaded area, which are evident during initial stages of wear, can be more accurately assessed. Also the service life performance of the prostheses joint system can be evaluated with greater accuracy and functional significance.

5. REFERENCES

1. R.M.Hall, A.Unsworth, P.Siney and B.M.Wroblewski, The surface-topography of retrieved femoral heads, *Journal of Materials Science-Materials in Medicine*, 7(1996) 739-744
2. A.Wang, C.Stark and J.H.Dumbleton, Role of cyclic plastic-deformation in the wear of UHMWPE actabular cups, *Journal of Biomedical Materials Research*, 29(1995) 619-626
3. J.Fisher, D.Dowson, H. Hamdzah and H.L.Lee, The effect of sliding velocity on the friction and wear of UHMWPE for use in total artificial joints, *Wear*, 175 (1994) 219-225
4. J.Zhao and E.T.Brown, Hydro-thermo-mechanical properties of joints in the carmenllis granite, *Quarterly Journal of Engineering Geology*, 25(1992) 279-290
5. T.W.Bauer, S.K.Taylor, M.Jiang and S.V.Medendorp, An indirect comparison of 3rd-body wear in retrieved hydroxyapatite-coated, porous, and cemented femoral components, *Clinical orthopaedics and related research* 298(1994) 11-18
6. B.Derbyshire, J.Fisher, D.Dowson, C.Hardaker and K.Brummitt, Comparative study of the wear of UHMWPE with zirconia ceramic and stainless steel femoral heads in artificial hip joints, *Med. Eng. Phys.* 16 (1994) 229-236
7. D.J.R.Cooper and J.Fisher, The effect of transfare film and surface roughness on lubricated wear of UHMWPE, *Clin. Matls.* 14 (1993) 295-302
8. I. Daubechies, *The lectures on wavelets*, SIAM, Philadelphia, PA, 1992
9. G. Strang and T. Nguyen, *Wavelets and filter banks*, Wellesley-Cambridge Press. 1996
10. X.Q.Jiang, L.A.Blunt and K.J.Stout, Comprehensive study on characterisation surface topography of orthopaedic joint prostheses using wavelets, *to be published in Journal of Inst. Mech. Eng.*

A New Reference Line to Surface Roughness Evaluation¹

CHEN Qinghu JIANG Xiangqian YANG Shunian LI Zhu

Research Centre of Measurement Technology & Scientific Instruments
Huazhong University of science and Technology
Wuhan, Hubei, 430074 P. R. China

1. The classical reference lines and its shortcomings

To find out an appropriate reference line is a key link of surface roughness evaluation. Generally, there are three kinds of classical reference line. One is the least squares mean line of a profile, which is a broken line. Strictly speaking, the reference line should be composed of some low frequency components characterizing the profile, including the form and the form error as well as the waveness of the profile. Therefore the reference line should be a smooth curve instead of a broken line. Besides, the least squares mean line is not unique certainly, it will be varied with variation of the sampling length or the sampling location. Another is the arithmetical mean line of a profile, whose shortcomings are the same as those of above one. Third is the multinomial regression line of a profile, which is described by a special algebraic expression. The special algebraic expression can only approximate to the low frequency components, its efficiency depends on the order of multinomial.

A new reference line named the wavelet reference line is proposed by the authors in this paper.

2. The mathematical model of wavelet reference line and its realization

Let $f(t)$ be a measured profile curve of the surface,
suppose

$$f(t) = s_1(t) + s_2(t) \quad (1)$$

$s_1(t)$ and $s_2(t)$ are satisfied with

$$S_1(w) = \begin{cases} F(w) & |w| \leq w_0 \\ 0 & \text{other} \end{cases} \quad (2)$$

¹ The project is supported by National Natural Science Foundation of China (No:59375255,)

$$S_2(w) = \begin{cases} F(w) & w_0 \leq |w| \\ 0 & \text{other} \end{cases} \quad (3)$$

$S_i(w)$ is Fourier transform of $s_i(t)$ ($i=1,2$). $F(w)$ is Fourier transform of $f(t)$.

We call w_0 as frequency boundary between the low frequency components and the high frequency components. If $s_1(t)$ includes the form and the form error as well as the waveness of surface profile, then $s_1(t)$ can be used as the reference line to surface roughness evaluatoin and $s_2(t)$ will be the description of surface roughness.

We use wavelet to realize formula (1). By using wavelet, we can decompose $f(t)$ as fallows:

$$f = f_{-1} + s_{-1}, \quad f_{-1} = f_{-2} + s_{-2}, \quad f_{-2} = f_{-3} + s_{-3}, \quad \dots, \quad f_{-N+1} = f_{-N} + s_{-N} \quad (4)$$

f_{-1} and s_{-i} are calculated by Multiresolution Analysis:

Let

$$\begin{cases} \Phi(t) = \sum h_n \Phi(2t - n) \\ \Psi(t) = \sum g_n \Phi(2t - n) \end{cases} \quad (5)$$

$\{C_n^0\}$ are satisfied with

$$f(t) = \sum C_n^0 \Phi(t - n) \quad (6)$$

and $\Psi_{-k,n} = 2^{-k/2} \Psi(2^k t - n), \quad \Phi_{-k,n} = 2^{-k/2} \Phi(2^k t - n),$

$$\begin{cases} f_{-k} = \sum C_n^k \Phi_{-k,n} \\ s_{-k} = \sum d_n^k \Psi_{-k,n} \end{cases} \quad (k=1,2,\dots,N) \quad (7)$$

$$\begin{cases} C_n^k = \frac{1}{\sqrt{2}} \sum_{j \in \mathbb{Z}} C_j^{k-1} \bar{h}_{j-2n} \\ d_n^k = \frac{1}{\sqrt{2}} \sum_{j \in \mathbb{Z}} C_j^{k-1} \bar{g}_{j-2n} \end{cases} \quad (k=1,2,\dots,N) \quad (8)$$

That is

$$f = s_{-1} + s_{-2} + \dots + s_{-N} + f_{-N} \quad (9)$$

s_{-i} ($i=1,2,\dots,N$) is the high frequency component of surface profile. f_{-N} is the low frequency component of surface profile.

Let $f(x)$ be discrete data, T sampling distance, and $\Omega = \frac{\pi}{T}$, selecting Shannon wavelet, we can prove

$$S_k(w) = \begin{cases} F(w) & \frac{\Omega}{2^k} \leq |w| < \frac{\Omega}{2^{k-1}} \\ 0 & \text{other} \end{cases} \quad (10)$$

That is , when $w_0 = \frac{\Omega}{2^N}$, let $s_2(x) = s_{-1} + s_{-2} + \dots + s_{-N}$, $s_1(x) = f_{-N}$, we can realize model (1).

3. Examples

Example 1: Figure 1 shows the measured curve of a workpiece profile. we get the wavelet reference line shown in Fig.2 and roughnes shown in Fig.3 from the curve shown in Fig.1 by using wavelet.



Fig. 1 Measured curve of workpiece profile



Fig. 2 Wavelet reference line of the curve Fig.1



Fig. 3 Roughness of the curve Fig.1

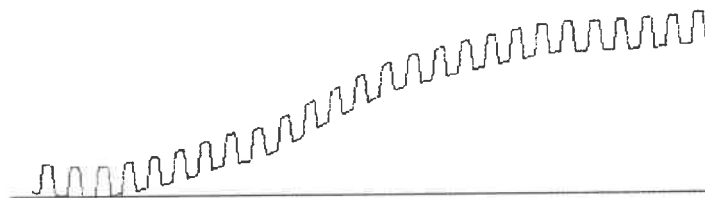


Fig. 4 The measured curve of Talysurf-5 roughness sample plate (No. 14502)

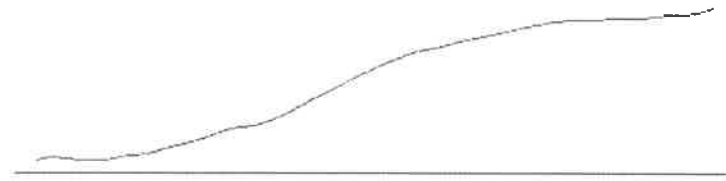


Fig. 5 The reference line to roughness evaluation

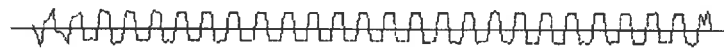


Fig. 6 The real roughness of the sample plate

Example 2: Figure 4 shows a profile curve of Talysurf-5 roughness sample plate (No. 14502), which is measured by a new instrument. Because of displacement error of the instrument, the measured curve is mixed with form error of the sample plate profile. By using Daubechies wavelet, we get an ideal reference line (as shown in figure 5). Figure 6 shows real roughness of the sample plate.

4. Advantages of wavelet reference line

From the examples mentioned above, it is obvious that the wavelet reference line has some advantages over the classical reference lines. Firstly, the wavelet reference line is a natural and smooth line, without algebraic expression. Secondly, it is unique and independent of sampling length and sampling location. Thirdly, it can get higher precision of surface roughness evaluation than that of the classical reference lines. Fourthly, it is a smooth arithmetical mean line of the profile, because of $\int s_2(x)dx=0$.

References

1. XIAO Shao-Jun, JANG Xiang-qian, XIE Tie-Bang, New Reference Line for Estimating the Roughness of the Arbitrary Curved surface, Proceeding of Measurement Technology & Intelligent Instruments, SPIE 1993, 2101(2):909-913
2. CHEN Qinghu, ZHOU Liping, XIE Tiebang, LI Zhu, A Method of Separating the Synthetical Topography of Surface with Wavelets, Journal of Huazhong University of Science and Technology (in chinese), No.5, 1997
3. I. Daubechies, Ten Lecture on Wavelets, CBMS-NSF Series on Applied Mathematics, SIAM, to appear, 1992

An Investigation of Filtering of 3D Surface Texture Measurements Using Scale-Sensitive Fractal Analysis and PSD

Donald K. Cohen

Michigan Metrology, Livonia, MI

Christopher A. Brown, William A. Johnsen, Pierre Hoch
Worcester Polytechnic Institute, Worcester, MA

1. Introduction

In addition to developments of new technology (Caber, 1993; Poon et al., 1995) used to measure the 3D surface texture of critical components, various 3D filtering techniques have been discussed and implemented in the various systems. Earlier work (Stout et al., 1993) proposed a number of filtering techniques for 3D surface data based on research using a 3D stylus-based contacting device. Recently Cohen et al (1992) described a device which was a commercialized, 3D, non-contact, texture measurement technique, referred to as Vertical Scanning Interferometry (VSI), which included various 3D filtering schemes. This paper will discuss the characteristics of the various filters implemented with the VSI technique via scale sensitive fractal analyses (Brown et al., 1993), power spectral analysis, and traditional texture parameters.

2. Scale-Sensitive, Fractal Analysis by the Patchwork Method

The patchwork method is a true 3D analysis that uses multiple virtual tiling exercises with triangular patches to determine area scale relations for a topographic data set (Brown et al., 1993). The patches represent the scale of measure, or observation. For an individual tiling exercise, the patches are kept the same size in 3D, although the shape varies slightly to allow complete coverage. Tilings are made over a range of scales, from the largest (about half the nominal or projected area of the data set) to the smallest (one half the square of the sampling interval). Fractal parameters are determined from a log-log, area-scale plot of the relative area versus the patch area. The relative area is the measured area (number of triangles used in a tiling exercise times the area of the triangles) divided by the nominal or projected area covered by the tiling. In this way the relative areas are related to the slopes on the surface as a function of scale (the reciprocal of the cosine of the angle between the normals to the patch and the nominal surface). The relative areas also indicate how much surface is available for interacting, at a particular scale.

In this study, the 1.5 decade region of the area-scale plot with the steepest negative slope is used to determine the area scale fractal complexity ($Asfc = -1000(\text{slope})$). The $Asfc$ is used for convenience and is related to the fractal dimension, D ($Asfc = 1000(D-2)$). The smooth-rough crossover (SRC) is the scale where the relative areas become significantly larger than 1 (in this case a relative area threshold of 1.00052 was used). At scales larger than the SRC, the surface appears smooth. At scales below, it appears rough. The area scale relations are of interest because they are related to the fractal dimension and they have a clear physical interpretation (Brown, 1994).

3. Filter Types

The VSI system used in this work was developed by the WYKO corporation (WYKO, 1996), and includes a number of 3D filtering techniques. The primary filtering techniques address the removal of any residual tilt, curvature, or cylindrical form included in the measurement from system alignment or sample geometry. The removal is performed by subtracting the best-fit "least squares" linear/quadratic terms. Once the surface form is addressed, the next level of filtering involves the resulting surface texture.

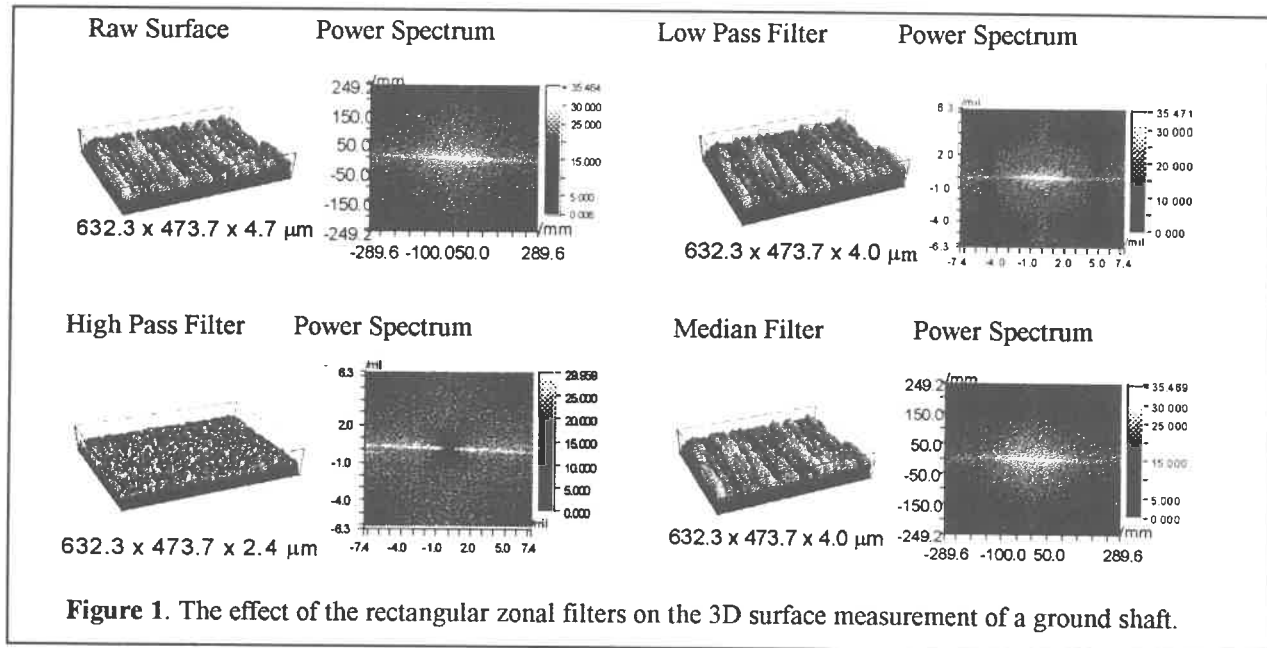
3.1 Convolution Filters

One class of filters based on convoluting various arrays through the 3D data set produces a new ‘filtered’ data array. The three convolution filters discussed here include a high pass, a median, and a low pass filter as shown below.

$H(p_1), H(p_2), H(p_3)$	High Pass Filter	$H(pc\text{-new}) = (8H(pc) - H(p_1) - H(p_2) - H(p_3) - H(p_4) - H(p_6) - H(p_7) - H(p_8) - H(p_9))/9$
$H(p_4), H(p_6), H(p_9)$	Median Filter	$H(pc\text{-new}) = \text{Median}(H(p_1), H(p_2), H(p_3), H(p_4), H(p_6), H(p_7), H(p_8), H(p_9))$
$H(p_7), H(p_8), H(p_9)$	Low Pass Filter	$H(pc\text{-new}) = (H(p_1) + H(p_2) + H(p_3) + H(p_4) + H(p_6) + H(p_7) + H(p_8) + H(p_9))/9$

$H(p_i)$ represents the height measured at the i th pixel and $H(pc\text{-new})$ is the new calculated height at the central pixel. The resulting new values are then placed in an additional array termed the “filtered” array.

The effects of the various convolution filters are illustrated on the surface of a ground shaft in Figure 1. We find that the low pass filter actually acts like a notch filter on a rectangular grid at about 2/3 of the surface bandwidth. Thus the low pass filter severely attenuates high spatial frequencies over a narrow band as opposed to over a gradual region. The high pass convolution filter appears to cleanly eliminate just the central portion (i.e. low frequency) of the spatial components within a rectangular region. Finally, the median filter appears to more uniformly accentuate the low frequency components without introducing major artifacts in the frequency domain.



Typically, amplitude parameters are dominated by lower spatial frequency components, as evidenced by the great effect the high pass filter has on the measured values (Boudreau, 1995). We find that for the ground surface, the R_a is only reduced by about 10% for the low pass and median filters.

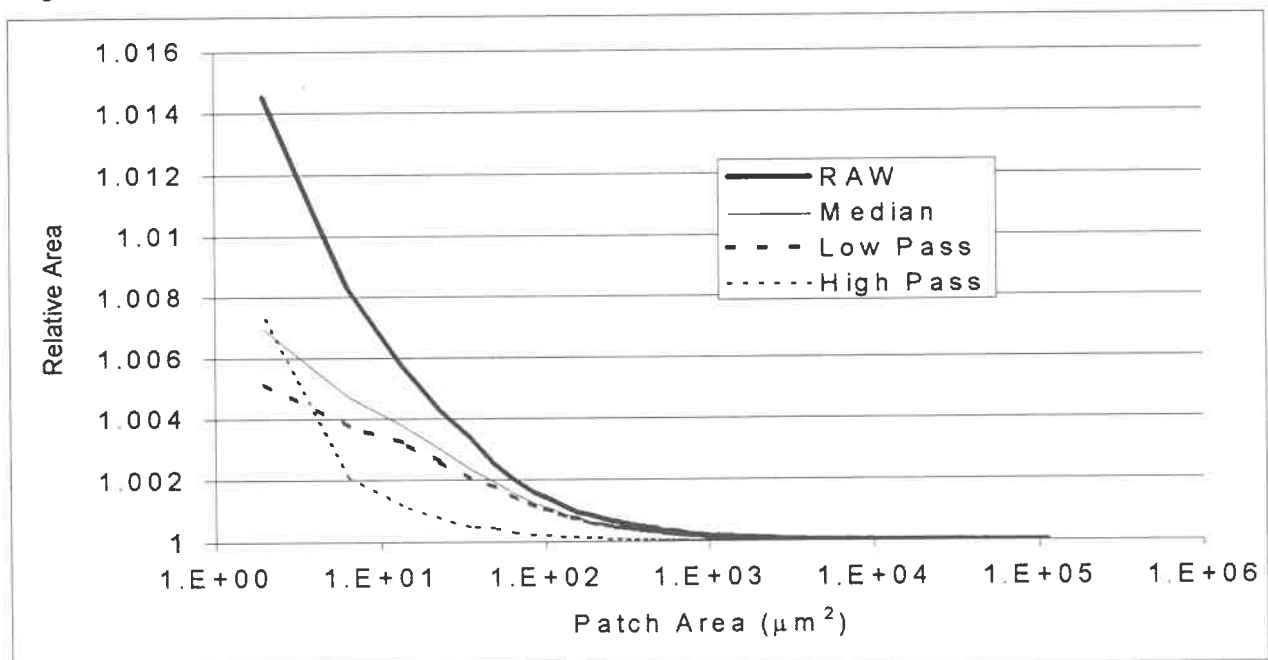
Table 1. Ground Shaft

	<i>parameters:</i>	conventional				fractal	
		R_a (μm)	R_z (μm)	X Slope R_q (deg)	Y Slope R_q (deg)	Asfc	SRC (μm)
	raw	0.41	3.71	10.4	5.20	3.66	360
<i>filter:</i>	high pass	0.12	1.83	7.8	4.2	2.05	32
	median	0.38	3.29	7.0	3.08	1.55	249
	low pass	0.37	3.19	5.7	2.35	1.06	231

We find a vast difference between the filters when considering other hybrid parameters such as the X and Y slope Rq (i.e. the RMS of the slope across the surface in the X and Y directions). We find that the low pass filter has the most dramatic effect on the surface slopes while the high pass filter only reduces the surface slope values by about 20%. This data suggests that the high spatial frequencies dominate the calculation of slope for this particular instrument / sample combination.

The three filters discussed here offer different features depending on the desired output. In earlier work (Cohen et al., 1997), using a low pass filter prior to applying conventional stylus-based filters, good agreement with conventional stylus amplitude parameters was obtained but poor correlation resulted for slope related parameters. This discrepancy may possibly be understood based on the characteristics of the low pass filter described here and its effect on the surface slope parameters. Thus, future work is suggested, to investigate the efficiency of the median filter when considering correlation of this optical based technique with stylus-based instruments.

Figure 2. Area-Scale plot of topographic data acquired from a ground shaft and treated with convolution filters.



3.1.1 Area-Scale Relations and Fractal Parameters

The relative areas calculated from the raw data are always greater than those calculated from the filtered data at scales below the SRC (Fig. 2). The SRC and the Asfc are greatest for the raw data (Table 1.). In all four analyses the 1.5 decade region with the steepest negative slope is at the finest scales. The high pass filter has the largest influence on the SRC and the least on the Asfc. This is logical as it has attenuated the largest features that have the greatest influence on the SRC and left the finest, which have the greatest influence on the Asfc. The median and low pass filters similarly have less influence on the SRC. A comparison the Asfc shows the low pass filter has removed the greatest amount of the fine scale complexity, followed by the median filter. As expected from the definition of the relative area, the reciprocal of the cosine of the X slopes in Table 2 are numerically close to the relative areas at the finest scale, although X slope is a 2D parameter.

3.2 Digital Filters

The final set of filters available is based on digital techniques in the spatial frequency domain. The surface data array is transformed via an FFT into the spatial frequency domain. A rectangular zonal filter is applied to the transformed data either passing the high spatial frequencies (i.e. the digital high pass filter) or the low spatial frequencies (i.e. the digital low pass filter). The modified spatial frequency array is then inversely transformed to produce the filtered texture data. To demonstrate the action of the digital filters, a turned brake rotor surface with a prominent lay pattern was chosen. This surface had a dominant spatial frequency component at about 5mm^{-1} (along the x-axis) and had an asymmetric cutting pattern. As for the convolution-based filters, we find that the amplitude parameters are slightly reduced by the low pass digital filter, as opposed to the slope parameters which are greatly affected. Conversely, when a high pass digital filter is applied, the amplitude parameters are greatly reduced whereas the slope parameters are left marginally unaffected. As depicted in Figure 3, the digital filters modify frequency space differently than the convolution filters.

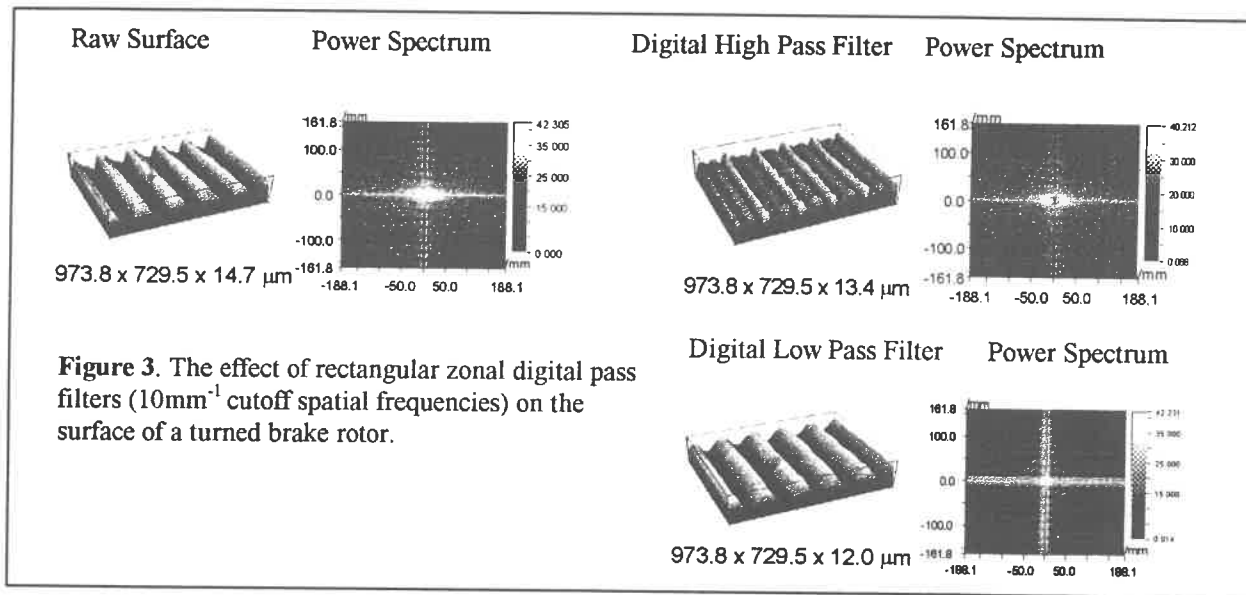


Figure 3. The effect of rectangular zonal digital pass filters (10mm^{-1} cutoff spatial frequencies) on the surface of a turned brake rotor.

Table 2. Brake Rotor

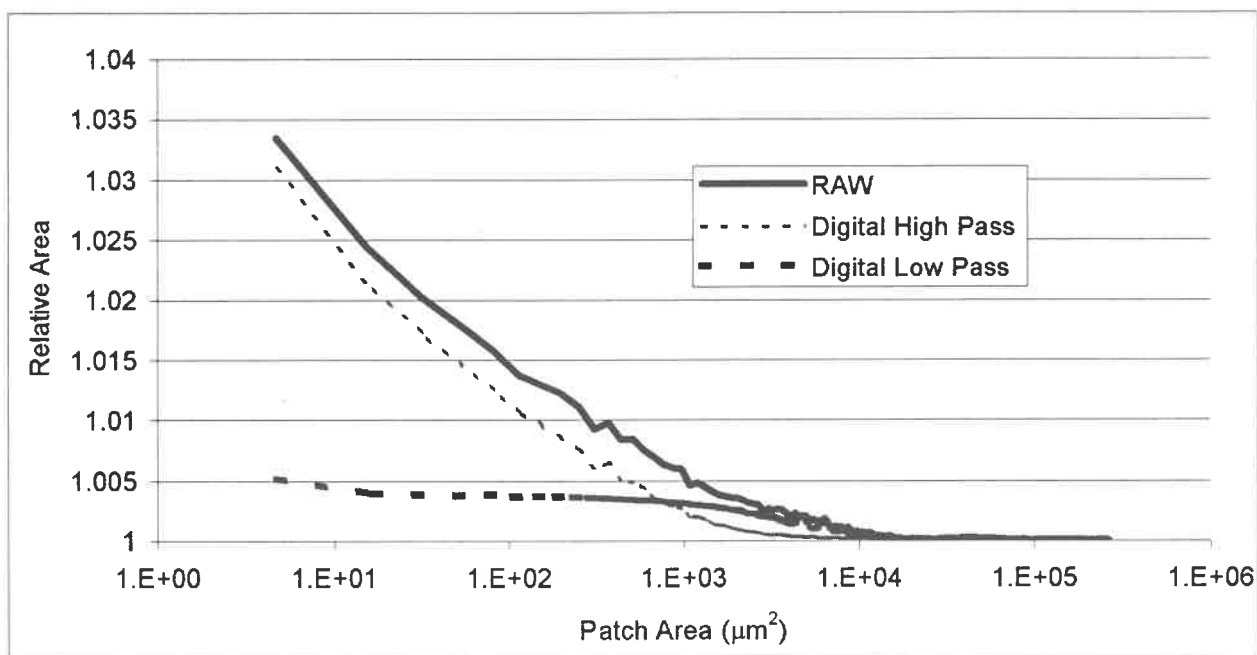
<i>parameters:</i>		conventional				fractal	
		Ra (μm)	Rz (μm)	X Slope Rq (deg.)	Y Slope Rq (deg.)	Asfc	SRC (μm)
	raw	2.73	12.08	16.6	7.86	5.91	12 146
<i>digital filter:</i>	low pass	2.27	10.32	4.8	2.0	1.11	9 432
	high pass	1.28	10.57	15.9	7.8	6.14	3 801

3.2.1. Area-Scale Relations and Fractal Parameters

Again the relative areas calculated from the raw data are always greater than those calculated from the filtered data at scales below the SRC (Fig. 4). The influence of the two digital filters on the area scale relations is clear. On Fig. 4 it appears that the digital low pass filter has removed almost all the features at scales below about $2000\mu\text{m}^2$ that contributed to increasing the relative area, and hence the complexity. The 1.5 decade region with the steepest negative slope on the low pass filter has a range of about 600 to $19,000\mu\text{m}^2$, whereas all the others are from 4.7 to $150\mu\text{m}^2$. The slope of the plot representing the low pass filtered data is near zero in the fine scale range, although it is close to the raw data in the range above $600\mu\text{m}^2$. The high pass filter appears to have shifted the area-scale plot for the raw data slightly less than a decade toward the finer scales. The effect of the shift is shown clearly in the lower SRC for the high pass filtered data, while the complexity increases slightly. The slight increase can be attributed to the removal of the features that contributed to the relative area of the scales at the high end of the 1.5

decade region, making the change in relative area toward the finer scales more abrupt. Again, the reciprocal of the cosine of the X slopes in Table 2 are numerically close to the relative areas at the finest scale.

Figure 4. Area-Scale plot of topographic data acquired from a turned brake rotor and treated with digital rectangular filters.



4. Summary

This work has been presented to alert the surface metrology community to the various 3D filtering techniques that are evolving in currently available instruments and gages. We demonstrate here that a number of issues exist as to which filter type and implementation should be considered for standardization in the future. This work demonstrates that among the filters presented, the median filter may offer the best compromise in suppressing high spatial frequency noise components while maintaining the nature of the surface.

The area-scale relations calculated by the scale-sensitive fractal analysis by the patchwork method provide a clear picture of the influence of the filtering. In general, the effect of filtering is to reduce the relative areas as a function of scale. The reduction at particular scales follows a logical progression based on that kind of filtering. There are clear differences between the digital and convolution filters, as can be seen from the area-scale plots.

5. References

- B.D. Boudreau, J. Raja, H. Sannareddy, A Comparative Study of Surface Texture Measurement Using White Light Scanning Interferometry and Contact Stylus Techniques, Proceedings from ASPE 1995 Annual Meeting, Vol. 12, 15-20 October 1995.
- C. A. Brown, "A Method for Concurrent Engineering Design of Chaotic Surface Topographies", Journal of Materials Processing Technology, 44 (1994) 337-344.
- C. A. Brown, P.D. Charles, W. A. Johnsen, S. Chesters, "Fractal Analysis of Topographic Data by the Patchwork Method", Wear 161 (1993) 61-67.

Paul J. Caber, "Interferometric profiler on rough surfaces," Appl. Opt. Vol. 32, No 19, 1 July 1993.

Donald Cohen, Paul J. Caber, Chris P. Brophy, Rough Surface Profiler and Method. United States Patent 5,133,601, 28 July 1992.

Donald Cohen, R.P. Garibay, J. McNeill, J. Huang, Correlation of Optical Surface Prolifometry with Contact Prolifometry on Precision Peened Surfaces, SME Conference on 3D Gaging and Technology, Detroit, MI, 4 March 1997.

Chin Y. Poon, Bharat Bhushan, Comparison of surface roughness measurement by stylus profiler, AFM and non-contact optical profiles, Wear 190 (1995) 76-88.

K.J. Stout, P.J. Sullivan, W.P. Dong, E.Mainsah, N. Luo, T. Mathia, H. Zahouni, The Development of Methods for the Characterization of Roughness in Three Dimensions. Publication No. EUR 15178 EN, ISBN 0 7044 13132, ECSC-EEC-EAEC, Brussels-Luxembourg and Authors, 1993.

WYKO NT-2000 Instrument Brochure, WYKO Corporation, Tucson, AZ 1996

A Fractal-Based Comparison of Surface Profiling Instrumentation

M. Malburg

Corporate Metrology, Cummins Engine Company, Inc.
Columbus, IN, USA

This paper explores the use of scale sensitive fractal analyses as a means of establishing the performance of instruments and for intercomparisons of instruments. The use of the compass or "coastline" method is employed in the comparison of profiles obtained from 11 common instruments/configurations. In addition to the compass method, Fourier analysis (which some argue is also a fractal approach) is employed in the comparison of the data sets. Finally, tip radius and filtering effects are explored in light of their influence on the fractal analyses.

Introduction

In many large companies it is not uncommon to find several different instruments employed in the measurement of surface profiles. This can become a problem when these instruments produce differing results upon measuring similar surfaces. This problem becomes very significant when measuring very fine surfaces where instrument resolution and frequency response become significant factors.

This paper presents the results from a study whereby fractal analyses were used as a means of distinguishing the performance of several surface measuring instruments.

Study Details

In order to ascertain the differences between instruments a somewhat "difficult" surface was selected - a relatively fine, plateaued surface. A profile from this surface is shown in Figure 1.

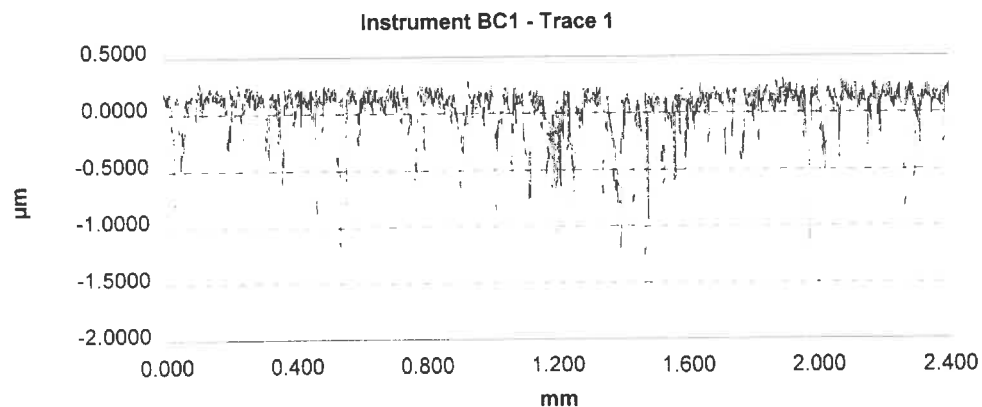


Figure 1 A profile from the test surface.

A rectangular area of the surface was indicated such that 10 equally spaced parallel traces 2.4 mm in length would be collected from each instrument. The test surface was of hardened tool steel thus the potential for damage during measurements was minimized.

Eleven measurement configurations were incorporated in the study based on different manufacturers, models and measurement settings. In all cases, the instruments were calibrated according to the manufacturer's recommendations. These configurations are "coded" based on the three character scheme. For example:

A B 3

A: Manufacturer
B: Model
3: Settings (e.g. traverse speed)

Fourier Analysis

Historically, Fourier analysis has been applied to determine frequency response characteristics of surface metrology instrumentation. However, in the analysis of these instruments, a Fourier transform of a single data set (the first profile) from each instrument yields a rather noisy "cloud" of data. (See Figure 2.)

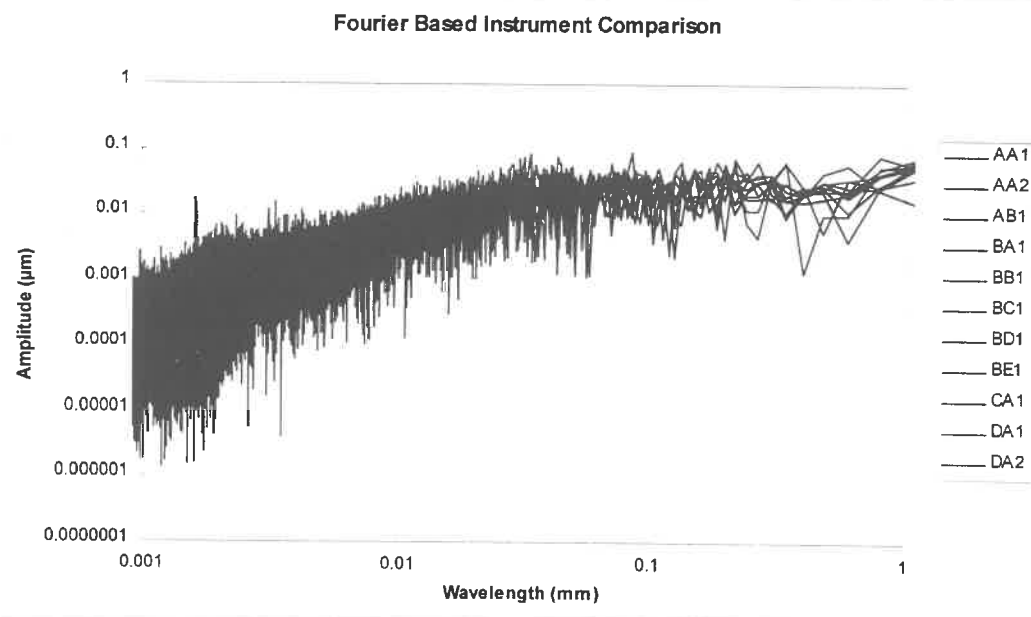


Figure 2 *Fourier analysis for 1 trace from each configuration.*

Some improvement in the readability of the Fourier results can be achieved through some averaging of the amplitude spectrum, however this process requires additional processing and there is still considerable overlap in the results. (See Figure 3.)

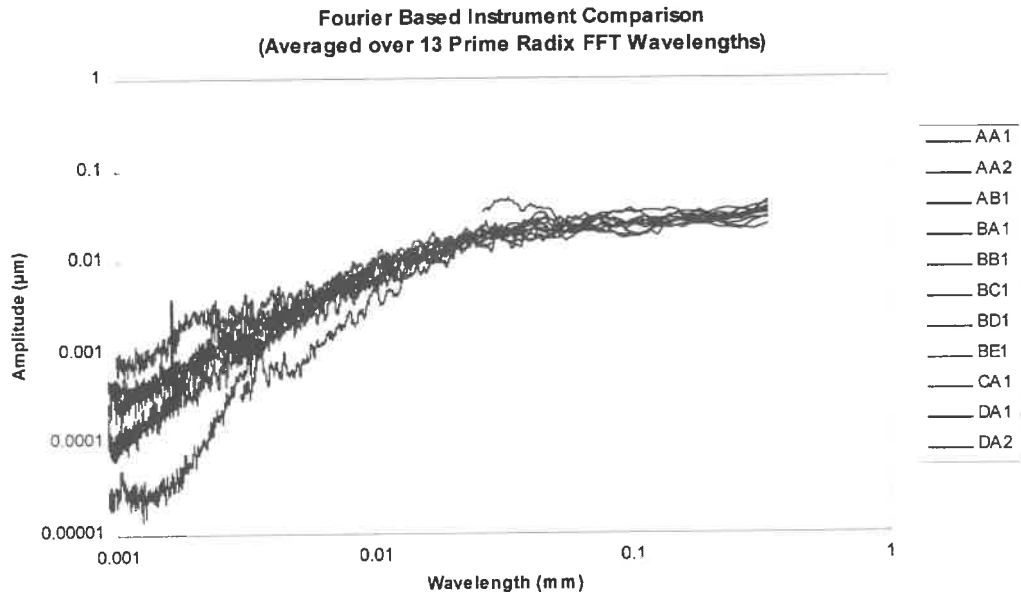


Figure 3 Averaging along the Fourier results.

Despite the apparent "noise", the Fourier approach has practical benefits in that this wavelength-based approach can be directly related to electronic or digital filtering as well as periodic environmental influences such as vibration or electronic noise.

Fractal Analysis

As was the case in the Fourier analysis, the first trace from each instrument was selected for analysis. The compass method was applied and the results are plotted in Figure 4.

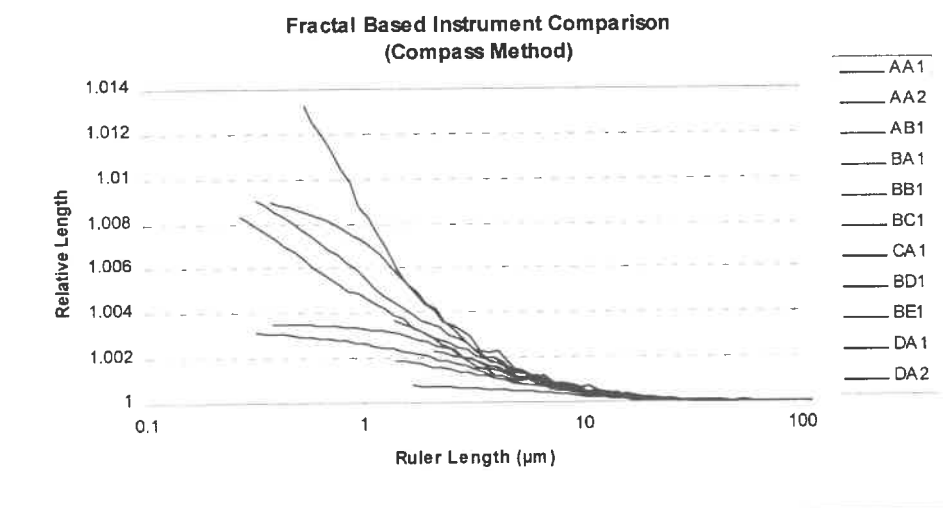


Figure 4 Fractal based instrument comparison.

The fractal analysis produced very interesting results as compared to the Fourier analysis. The most significant aspect of this approach is that there is no significant "noise" in each plot. Furthermore, each instrument is very distinctly represented in comparison to the others.

While Figure 4 demonstrates that the fractal-based approach can certainly distinguish between instruments. It still must be established whether this separation is related to the instrumentation or the differences in specific measurement location on the surface. One way to determine the influence of measurement location is to superimpose the fractal analyses from all ten traces from a single instrument. Configuration BC1 was chosen since it represented the highest "effective resolution" based on the compass method. The fractal analysis for all 10 traces from BC1 are presented in Figure 5.

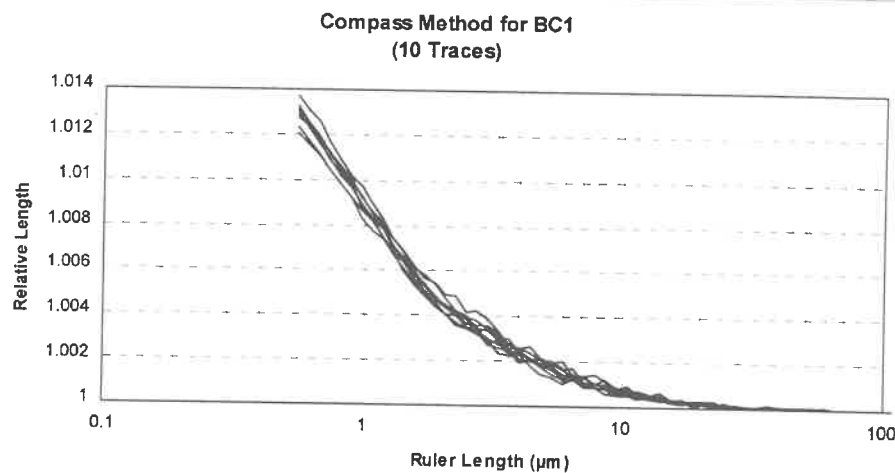


Figure 5 Analysis of 10 traces from the same instrument configuration.

The similarity in the Figure 5 results are very interesting in that the 10 traces were parallel and spaced over an area of a plateau surface with a cross-hatched lay. This cross-hatched property of the surface ensures that no two parallel traces will be the same (as shown in comparing the Figure 6 profile below with the Figure 1 profile above).

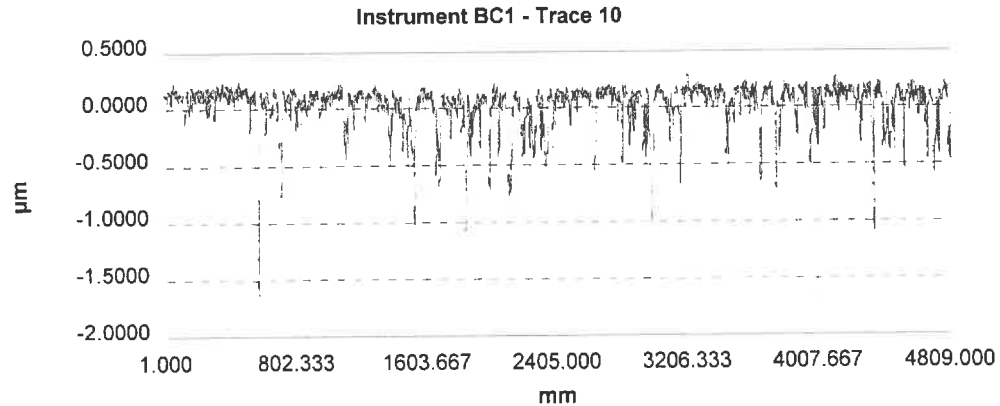


Figure 6 Profile #10 of the test surface for comparison with Figure 1 (profile #1).

Given this difference in profiles (Figures 1 & 6), the similarity in the fractal analyses (Figure 5) seems very significant. This may indicate that the fractal analysis (particularly when computing the slope of the line or "complexity") is more based on the general *character* of the surface rather than the specific positioning of the *details* of the surface.

Influences on Fractal Analyses

It has been shown above that the fractal analysis seems relatively consistent within a configuration. Thus, it becomes necessary to establish which aspects of the measuring process are responsible for the differences shown in Figure 4.

One aspect of the measuring process which varies from instrument to instrument is the application of analog or digital filtering. As an example of the influence of a digital filter, a 2.5 µm Gaussian (long-pass) filter was applied to trace #1 from instrument configuration BC1. The results are shown in Figure 7.

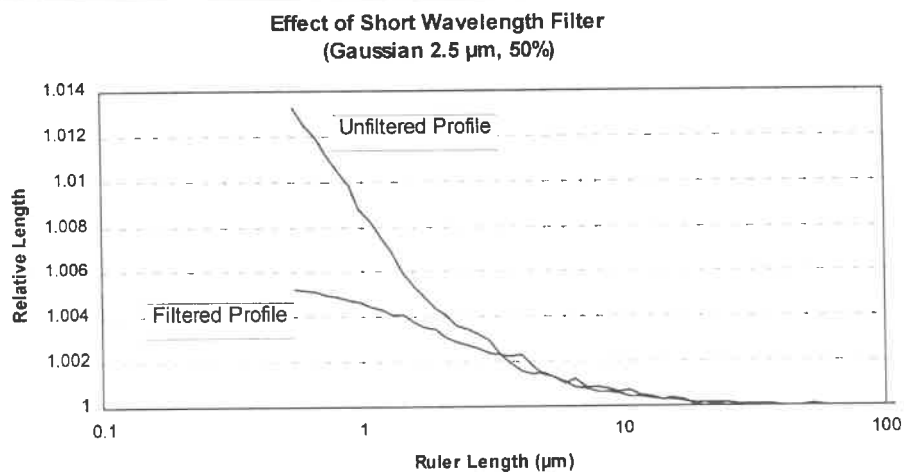


Figure 7 The influence of a long-pass, short wavelength filter.

The stylus tip radius is another aspect of the measuring process that can effect the instrument's ability to transmit short wavelengths or resolve small features. To ascertain the influence of tip radius variation on the fractal analysis, the BC1 trace #1 data set (as measured with a nominally 2 μm tip radius) was used in the numerical convolution of a 5 μm tip radius. The results are plotted in Figure 8.

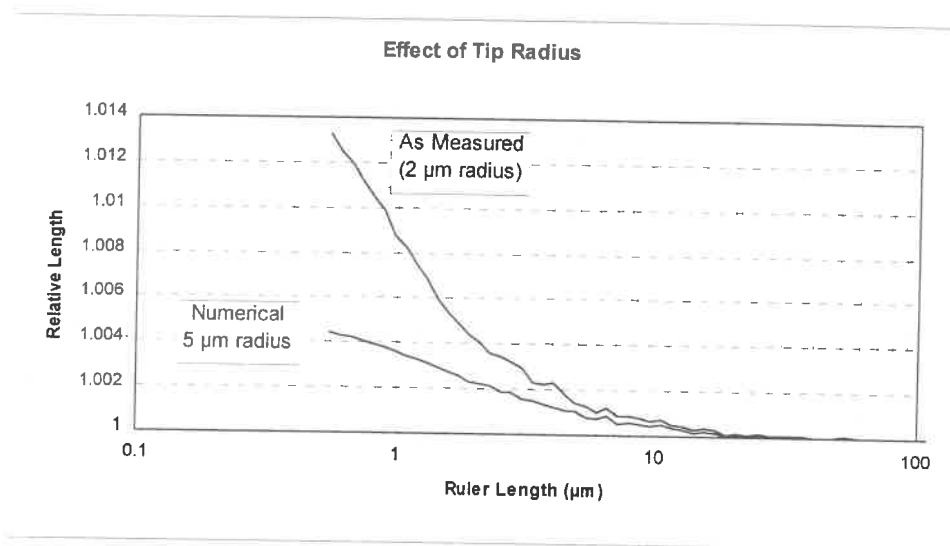


Figure 8 *The influence of tip radius.*

The similarity in results from Figures 7 and 8 are indicative of one of the problems with applying fractal analyses in the comparison of instruments. Specifically, fractal analyses may not be as "diagnostic" as Fourier analysis in the traditional sense. In the fractal based analyses, the influence of tip radius is shown to be very similar to that of the filter. Thus, when addressing the cause for differences between instruments, other more specific testing may need to be conducted.

Conclusion

Scale-sensitive fractal analyses have been shown to be useful in the comparison of surface profiling instruments. This approach seems to generate very reproducible data and which is relatively free from noise as compared to Fourier methods. The drawback of this approach is that it may not be as diagnostic in determining "why" differences exist between instruments as was demonstrated through the application of filters or varying tip radii.

Nonetheless, the emerging field of scale-sensitive fractal analysis appears to provide some very interesting tools which can certainly benefit the surface metrology community and much more work in the development, application and standardization of these methods is certainly warranted.

Acknowledgments

The author wishes to extend his appreciation to Dr. Chris Brown at Worcester Polytechnic Institute and Dr. Jay Raja at the University of North Carolina at Charlotte for the collaboration in this work.

A New Comprehensive Measuring and Analyzing System for Form and Surface Topography*

XIAO Shaojun JIANG Xiangqian XIE Tiebang LI Zhu

Research Centre for Measurement Technology and Scientific Instruments
Huazhong University of Science and Technology
Wuhan, Hubei, 430074 P.R.China

ABSTRACT

The sensor for measuring the form and surface topography, especially for precision curved surface, must have the character of wider dynamical measuring range and higher resolution in the meantime. A small-sized stylus-grating sensor, which using a Reflection Cylindrical Hologram Diffraction (RCHD) grating and a laser diode, is improved on the authors' previous researches. The sensor's measuring range is 6mm, and its resolution is 1nm. The sensor is used successfully in the comprehensive measuring and analyzing system for form and surface topography.

1. Introduction

The high precision measurement for form and surface topography is very important in the fields of mechanical, electronic and optical, etc.. Many sensors are successfully used to measure the form errors and the surface roughness separately. Generally, to measure the form errors, the sensor has wider dynamical measuring range with lower resolution; to measure the surface roughness, the sensor has higher resolution with less dynamical measuring range. For measuring the form errors and surface roughness in the meanwhile, for measuring the precision curved surface as an example, the sensor must have the character of wider dynamical measuring range and higher resolution in the meantime.

The sensors based on laser interferometer have become powerful tool in higher precision measurement. But the series of these sensors only have higher resolution (from 0.1nm to 1nm), and have not enough wider dynamic measuring range (less than 500 μ m). So that these sensors have not enough character mentioned above.

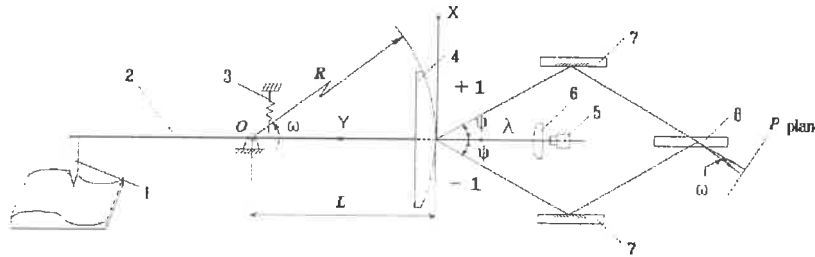
A small-sized stylus-grating sensor, which using a Reflection Cylindrical Hologram Diffraction (RCHD) grating and a laser diode, is improved on the authors' previous researches[1,2]. By the cylindrical surface of grating, the rising-falling movement of the stylus on the measured surface are spread to a motion as same as the linear motion of a plane grating. So that, the sensor's dynamic measuring range is mainly determined by the effective circular length of the RCHD grating; the sensor's resolution is mainly determined by the dividing precision of the grating's interference fringes. A laser diode is adopted as the interference light source instead of a He-Ne laser. The astigmatic property of the laser diode is the main obstacle to practical application. Especially for the grating interference, the collimating effect of laser diode is more strict than that of Michelson interference.

* Supported by National Natural Science Foundation of China (NSFC) (No. 59375255, 59575083)

2. Principle of the system

The comprehensive measuring and analyzing system for form and surface topography is mainly composed of four components: the stylus-grating sensor, the circuit for signals processing, X-Y two-direction precision moving table, and a computer with relevant controlling-measuring software[1,2]. The stylus-grating sensor, which determined the measuring range and resolution of the system, is the key component of the system, and its optical principle is introduced as follows.

Based on Morie interference theory, the grating interference means that two beams of light, which are diffracted on the grating and have different frequency between each other, form heterodyne interference which producing alternate dark and bright fringes. Its optical principle is shown as Fig. 1.



1. Stylus; 2. Level; 3. Spring; 4. RCHD grating; 5. Laser diode;
6. Pillar prism; 7. Reflecting mirror; 8. Refraction lens.

Fig.1 Optical principle of the stylus-grating sensor

The circular center of RCHD grating surface coincides with its rotation center O. Whenever the RCHD grating rotates around the point O, the diffraction angle ψ always keeps constant. So the cylindrical surface of the RCHD grating can be spread along its tangent direction, the practical optical principle of the RCHD grating is as same as that of the plane grating. When the plane light wave with the wavelength λ incidents to the RCHD grating along Y axis, it produces Fraunhofer far field diffraction. Based on the Doppler effect of light, the frequency of the order q diffracting light is:

$$f_q = f_0 + (v / \lambda) \sin \psi_q \quad (1)$$

where, v is the tangent speed of grating rotating; ψ_q is the diffraction angle of order q lightwave. Substituting the grating equation $d \sin \psi_q = q\lambda$ into Eq.1, then

$$f_q = f_0 + (v / d)q = f_0 + \Delta f_q \quad (2)$$

$$\Delta f_q = (v / d)q \quad (3)$$

where, Δf_q is the Doppler frequency displacement of order q lightwave; d is the grating constant. The positive and negative first orders of diffraction are adopted as the interference lightwave. Their Doppler frequencies are as follows:

$$f_{+1} = f_0 + (v / d) \quad \text{and} \quad f_{-1} = f_0 - (v / d).$$

The above two light beams make heterodyne interference. The beat frequency is:

$$\Delta f = f_{+1} - f_{-1} = 2v / d \quad (4)$$

The light intensity received by the PIN detector on P plane is:

$$\begin{aligned} I(t) &= I_1 + I_2 \cos(2\pi \int_0^t \Delta f dt) = I_1 + I_2 \cos(4\pi / d \int_0^t v dt) \\ &= I_1 + I_2 \cos(4\pi x / d) \end{aligned} \quad (5)$$

where, I_1, I_2 are the light intensity of the positive and negative first diffraction lightwave respectively; x is the displacement of grating on tangent direction. When the light intensity changes one whole period, that is the interferometer fringes moves one fringe interval, its phase changes 2π . So $2\pi = 4\pi x / d$, $x = d / 2$. This shows that the measuring signal only depends on the grating constant d rather than the wavelength λ since that x also responses to the displacement of stylus, and the pulse equivalent is equal to half of the grating constant. The interference fringes is divided and counted automatically by the system, so the information of form errors and surface roughness will be obtained. Through analyzing on theory and experiments, the sensor's resolution is about 1nm, the dynamic measuring range is 6mm.

3. Collimating of the Laser diode lightwave

Because the laser beam on the perpendicular and parallel directions appears from different source points on the junction plane of the laser diode (Fig.2), so the light beam appears as an ellipsoid on its cross section, but its wavefront is spherical. The astigmatic distance Δf is defined as the distance between the two apparent sources.

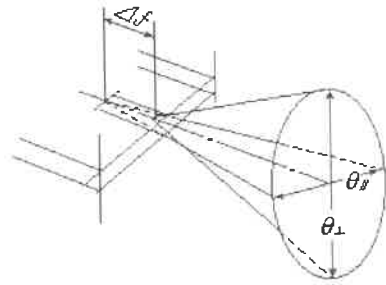


Fig.2 The astigmatic distance Δf of laser diode

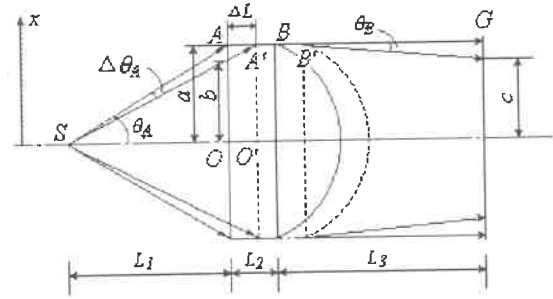


Fig.3 The collimating principle of pillar prism

As the interference light source, the laser beam must be collimated to a beam with parallel radiating, plane wavefront and even intensity in all directions. A pillar prism is adopted as the collimating lens, its incident end is a plane and the exit end is a spherical surface with radius R (Fig.3).

The traslation matrix of this pillar prism is: $[x_B, \theta_B]' = M * [x_A, \theta_A]'$

$$M = \begin{bmatrix} 1 & 0 \\ 0 & \frac{\eta_1}{\eta_2} \end{bmatrix} \begin{bmatrix} 1 & \frac{L_2}{\eta_2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{\eta_1 - \eta_2}{\eta_1 R} & \frac{\eta_2}{\eta_1} \end{bmatrix} \\ = \begin{bmatrix} 1 + \frac{\eta_1 - \eta_2}{\eta_1 R} \times \frac{L_2}{\eta_2} & \frac{L_2}{\eta_1} \\ \frac{\eta_1 - \eta_2}{\eta_2 R} & 1 \end{bmatrix} \quad (6)$$

where, η_1 is the refraction ratio of light beam in air; η_2 is the refraction ratio of light beam in prism; L_2 is the length of the pillar prism. Therefore,

$$x_B = (1 + \frac{\eta_1 - \eta_2}{\eta_1 R} \times \frac{L_2}{\eta_2}) x_A + \frac{L_2}{\eta_1} \theta_A \quad (7)$$

$$\theta_B = \frac{\eta_1 - \eta_2}{\eta_2 R} x_A + \theta_A \quad (8)$$

For calculating the average light intensity on plane G , the focus length of the light source point S and the incident light intensity must be calculated. The interval between the light source point S and its image plane G is not changed.

Because $L_2 \ll L_1 < L_3$, let $x_B = x_A$ anywhere. Suppose $\theta_B = 0$, when $SO = L_1$; Then $\theta_B = \Delta\theta_A$
 $= \arctg\left(\frac{a}{L_1}\right) - \arctg\left(\frac{a}{L_1 + \Delta L}\right)$, when the collimating prism is translated ΔL . The focus length on plane G is $2c$:

$$c = a - (L_3 - \Delta L) \times \tg\theta_B \quad (9)$$

The incident light intensity may be indicated by the length $2b$. So the average light intensity on plane G is:

$$\begin{aligned} \frac{2b}{2c} &= \frac{a \times L_1 / (L_1 + \Delta L)}{a - (L_3 - \Delta L) \times \tg\theta_B} \\ &= \frac{c(1)\Delta L + c(2)}{\Delta L^3 + c(3)\Delta L^2 + c(4)\Delta L + c(5)} \end{aligned} \quad (10)$$

where, $c(1)$, $c(2)$, $c(3)$, $c(4)$, $c(5)$ are constants depended on a , L_1 and L_3 .

The collimating effect of this pillar prism is the same in all directions. There is a alternate distance between the average light intensity curves of perpendicular and parallel directions respectively caused by the astigmatic distance Δf (Fig.4). The average light intensities of two orthogonal directions are equal to each other when the prism locates on the position B. So the laser beam is suitable for the grating interference.

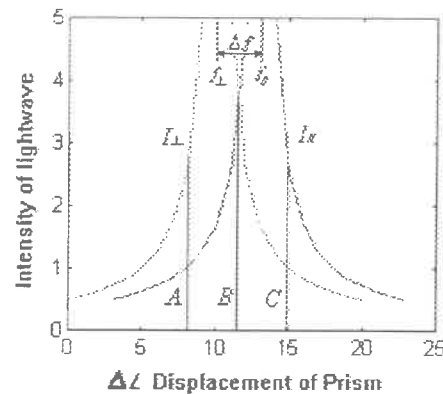


Fig.4 The curves of the average light intensity

4. Conclusions

Experiments shown, using the laser diode, the interference fringes are as good as using a He-Ne laser and the system is suitable to measure the form and surface topography.

References

1. JIANG Xiangqian, XIE Tiebang, YAO Caixian and LI Zhu, "Grating technology for topography measurement of curved surface". *SPIE*. Vol.2101. pp.554—557. 1993.
2. JIANG Xiangqian, GU Tingxi and LI Zhu, "Research on a new measuring and analysing system for curved surface topography". *SPIE*. Vol.2101. pp.1163—1167. 1993.
3. XIAO Shaojun, JIANG Xiangqian and XIE Tiebang, "New reference line for estimating the roughness of the arbitrary curved surface". *SPIE*. Vol.2101. pp.909—913. 1993.

Study on Polishing Mechanism of Magnetic Disc Substrate

Xiao Ming Liu (1), Guang Qi Cai (1) and Xiao Lin Feng (2)

(1.) Northeastern University, P. R. China

(2.) HEN LAN Electronic Equipment Inc., P. R. China

Abstract

In this paper, the process of polishing NiP plated magnetic disc substrate is introduced. The research aims to clear experimentally the influence of polishing pressure and speed on flatness, roughness and removal rate. Kinetic characteristics of grains and the micro-structure of surface are also investigated. The polished surface is analyzed with fractal geometry, which shows that the surface is fractal in structure.

Keywords: Polishing, Magnetic disc substrate, Polishing mechanism

1. Introduction

Recently, computer technology has tended micro-miniaturization and high density magnetic area. In order to increase memory density of magnetic disc, more excellent disc substrate obtained by ultra-precision machining is required [1]. Polishing operation plays an important role in improving roughness and accuracy of disc substrate. Many researchers such as Yasui and Sato [2], Mitsuhashi [3], Kawamura [4], Kasai, etc., paid more attention on the study of polishing mechanism of disc substrate. In addition, many studies have been made in those aspects which have close relationship to disc polishing [5,6]. The contributions have practical significance in improving disc substrate quality and machining efficiency. Based on the works above, this research will give an investigation on polishing mechanism in some aspects.

2. Experimental conditions

The experiments were carried out on a 6S-R4 double surface precision polishmaster made in Howard Strasbaugh Inc. in a precision machining shop. The experimental conditions are shown in Table 1. The Model MiniFiz-100 instrument with accuracy of 1/20 of laser wave length made in Phase Shift Technology Inc. and Dekak3ST Surface Profiler with

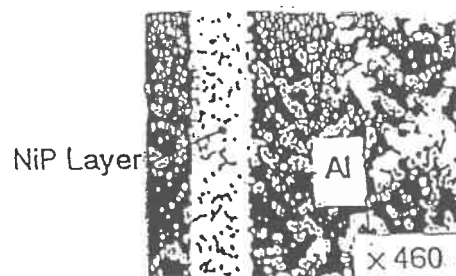


Fig. 1 Disc section

Table 1 Polishing conditions

Work-piece	Material	Plated NiP Aluminum Substrate (Fig. 2) 95 × 25 × 0.78 mm
	Size	
polishing Parameter	Carrier Speed	3-12 1/min
	Oscillating Speed	40-120 1/min
	Cylinder Pressure	70-140 kPa
	Eccentric	40 mm
Polishing Liquid	Abrasive	0.2 μm-Al ₂ O ₃
	PH	4.5-5
	Concentration	5 wt %-Al ₂ O ₃
	Pressure	280 kPa

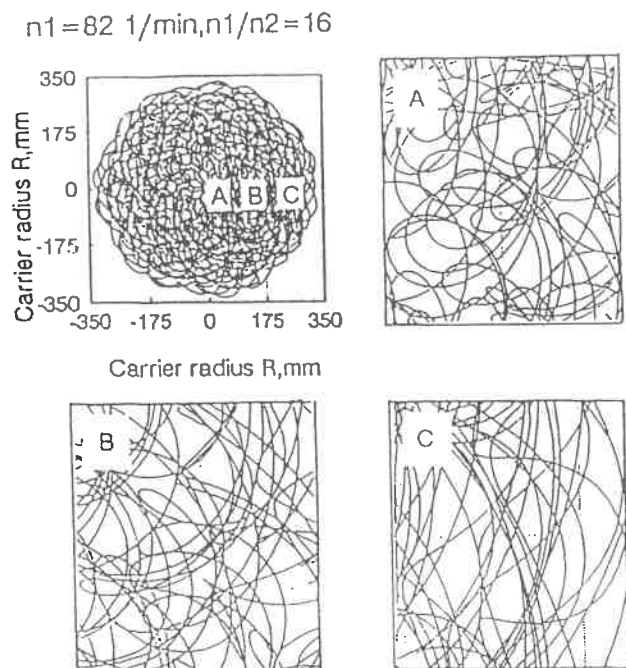


Fig. 2 Loci of 50 random grains on carrier

resolution of 0.1 nm made in Veeco Instruments Inc. were used to measure the flatness and roughness respectively. The thickness of plated layer was measured with Fischel Instrumentation. Measuring temperature was kept within $19 \pm 0.5 \text{ }^\circ\text{C}$.

3. Kinetics analysis

3.1 Loci of grains

Suppose the grains may embed onto polisher surface, the loci can be considered as the scratches of grains. Obviously, the loci of grains are different with their position on polisher. Only the influence of summing up of many random grains may have practical meaning. Fig. 2 shows the loci on carrier of 50 random grains which obey uniform distribution in two dimensions and its three enlargements at A, B and C. It can be seen that the loci composition is different, i.e., most line segments of loci at A are vertical to disc motion direction and those at C are even. The micro graphs and statistical results in Fig. 3 proved this also. The statistical analysis shows that the n_1/n_2 has influence on it. The uniformity can be improved when n_1/n_2 equals 14. In addition, the density of loci at different carrier

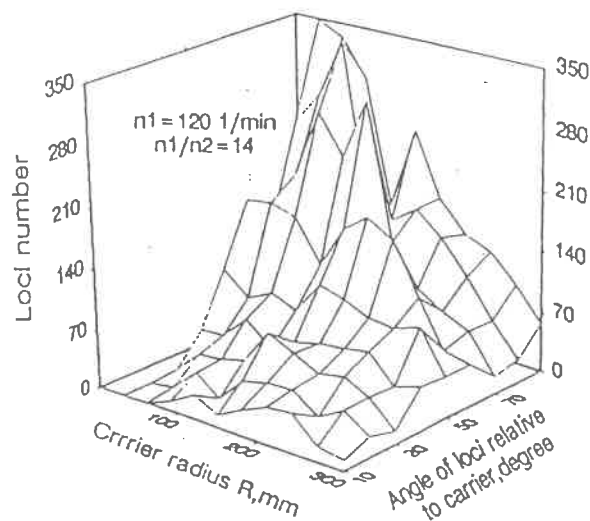


Fig. 3 Passing numbers of loci with different angles at different radius

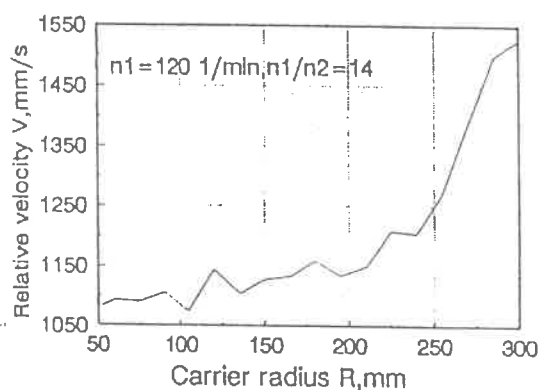


Fig. 4 Velocity distribution of a grain on carrier

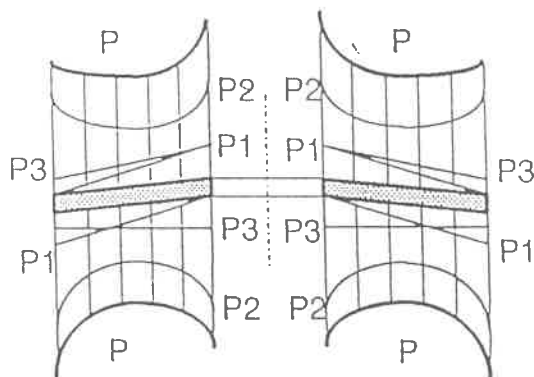


Fig. 5 Pressure distribution

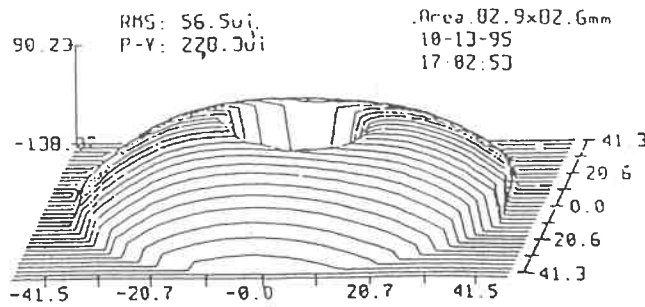


Fig. 10 Bowl error of disc substrate

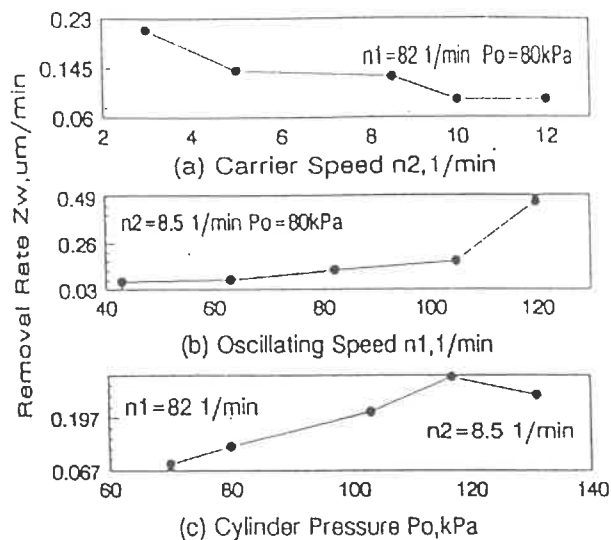


Fig. 11 Influence of n_1 , n_2 and P_0 on flatness

respectively (Fig. 5).

Beside those stated above, the following parameters also have influence on P: (1) Cylinder pressure; (2) Elasticity, thickness and primary flatness of polisher and disc substrate; (3) Pressure, and viscosity of polishing liquid; (4) Carrier speed; (5) Oscillating speed and Eccentric E of polishing plate. All of them must be paid more attention in industrial production.

As pressure is larger at periphery than at other position, the disc substrate may subside at its periphery as shown in Fig.5. The load F caused by polisher can be obtained by integral of the P . As

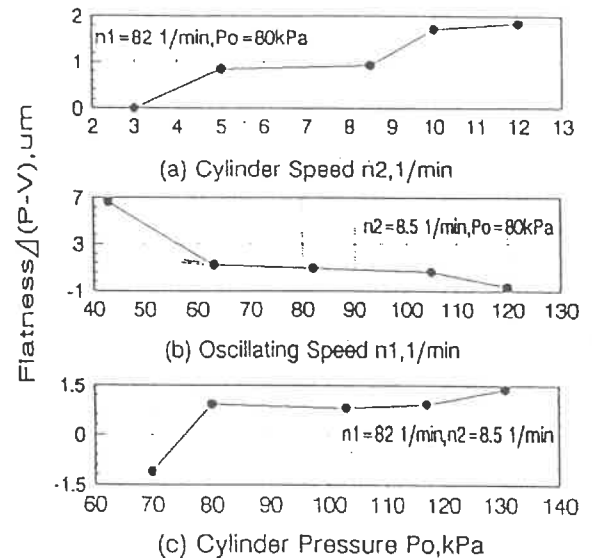


Fig. 12 Influence of n_1 , n_2 and P_0 on removal rate

the F is alternative in some range owing to the oscillating of polishing plate, the pressure on disc substrate is variable in polishing process.

4. Formation of polishing surface

Fig. 6 shows the micro-texture of plated NiP substrate surface taken by EPAM and SEM micrographs of polished surface. It can be seen that the primary surface (Fig. 6(a)) still remains lapping scratches in some extent although covered with NiP layer. It is these remained scratches and ruggedness of plated layer itself that form primary roughness. The roughness is reduced obviously after polishing (Fig. 6(b)). Except those fine scratches which is formed by single grains, there are many- scratches with depth of 1-3 nm and width about several nm (Fig. 6(c, d)). It can be explained that some of grains are bounded by chips to form some scraps of grain film and it is these scraps that scrape surface to form the wide scratches. In fact, the grains and scraps are mainly to plough disc substrate surface to cause plastic flow of material and swell at side of scratch. The chips are formed by extruding and flowing of material repeatedly and then peeling off for material fatigue. The veins are constructed by the overlapping of these scratches. As the compositions of surface and deeper layer are almost same by probing, It can be said that mechanical actions play stronger roles than chemical interactions in this polishing process. Material removal is mainly done by mechanical actions such as plastic deformation and / or brittle fracture. Consequently, the primary factors in this

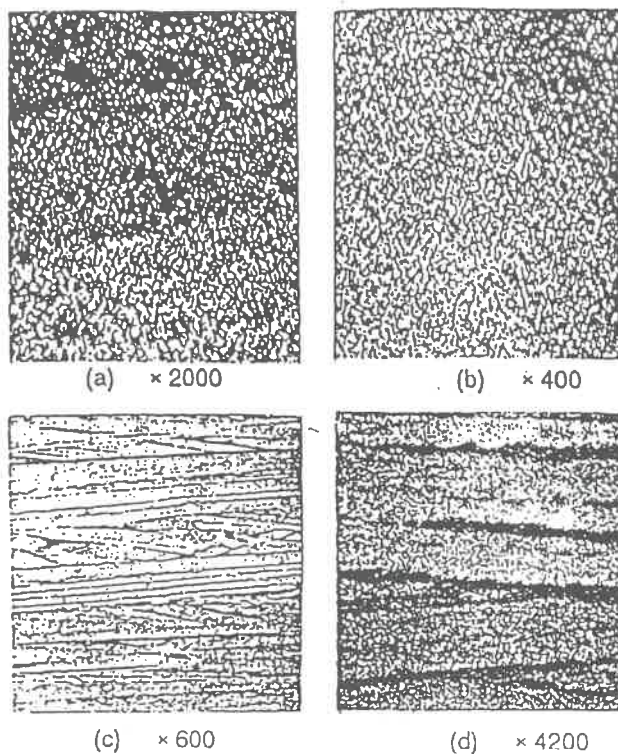


Fig. 6 Features of surface

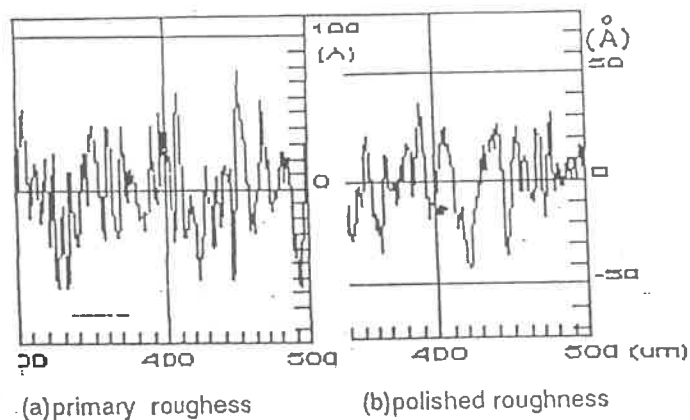


Fig. 7 Surface roughness profile

radius is also different and varies with n_1/n_2 . Statistical results shows that the density become more uniform when n_1/n_2 equals 14. In such case, the polished surface tends to be more isotropic.

3.2 Relative velocity of grains

As the relative velocity of each grains is variable with its position (Fig. 4), the mean velocity V of grains at different radial position is different, i.e., the V increase with the carrier radius. The n_1/n_2

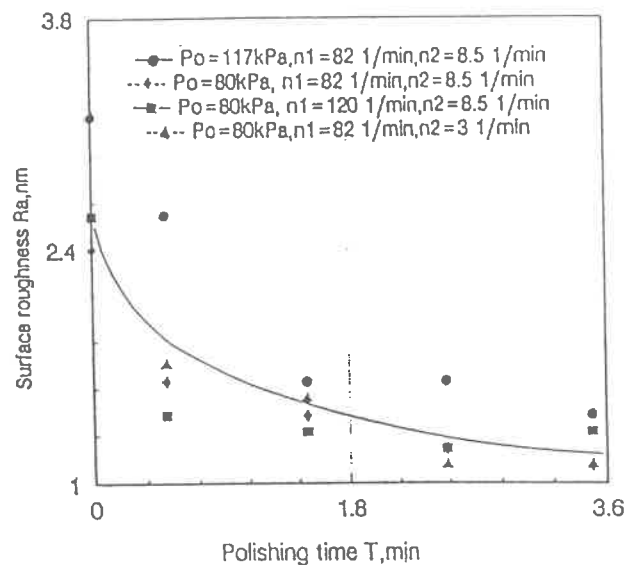


Fig. 8 Influence of polishing time on Ra

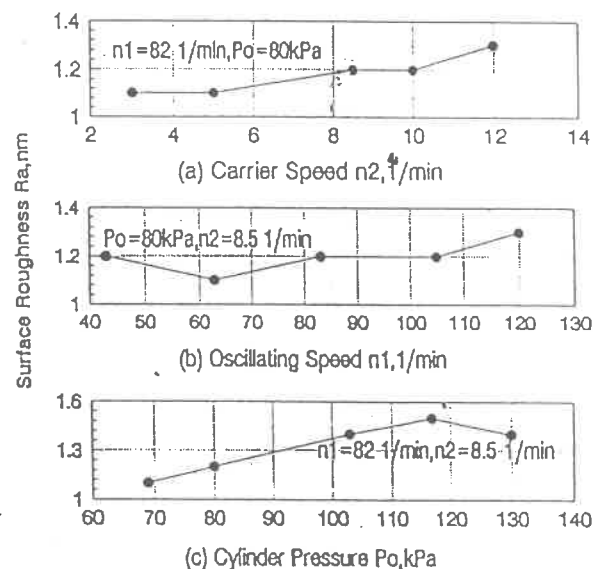


Fig. 9 Influence of n_1 , n_2 and P_0 on Ra

also has influence on V , and the variation of V is the least when n_1/n_2 equals 27.

3.3 Pressure distribution

In double-side polishing process, the actions between polyurethane polisher and disc substrate can be simplified as shown in Fig. 5. The pressure P on disc substrate consists of P_1 , P_2 and P_3 which are caused by bowl error of disc substrate, unbounded contact between polisher and disc substrate, and static and dynamic pressure of polishing liquid

fracture. Consequently, the primary factors in this polishing process include: (a) material of grains and polisher; (b) loci and velocity of grains and (c) pressure and time of polishing. This research aims at (b) and (c) when (a) is determined.

5. Experimental results

Fig. 8 shows the influence of polishing time on surface roughness at different conditions. The roughness R_a decreases noticeably at the beginning and keeps within 1-1.5 nm after 1.5 minutes. It can be explained that the pressure between grains and workpiece decreases greatly after initial polishing stage, and then the roughness is only influenced by the height of scratch swell.

Fig. 9 shows the influence of n_2 , n_1 , and cylinder pressure P_0 on surface roughness after three minutes polishing. It shows that n_1 , n_2 and P_0 all have little influence on roughness. The roughness of 1-1.5 nm can be obtained at all these experimental conditions.

Generally, the plated NiP disc substrates have bow error (Fig. 10). As the error is different to discs, in this research, the increment $\Delta(P-V)$ of the peak-valley height of substrate surface between primary substrate and polished one is taken to measure its flatness. Fig. 11 shows the influence of n_1 , n_2 and P_0 on the $\Delta(P-V)$. With the increment of n_1 and decrement of n_2 and P_0 the flatness is improved.

Fig. 12 shows the influence of n_1 , n_2 and P_0 on removal rate. The removal rate increases with the increment of n_1 and P_0 and decreases with n_2 . In Fig. 12 (c), the decrement of removal rate when P_0 passed over 117 kPa may be caused by the decrement of grains' active edges owing to grains embedding onto polisher deeper.

6. Fractal description of surface roughness

Many surfaces which seem smooth show extreme ruggedness such as the surface of polished disc substrate (Fig. 8). The surface profile appears random, multiscale and disordered. It has fractal characteristics which are continuous, non differentiable and statistically self-similarly. It can be characterized by fractal geometry to yield a parameter—fractal dimension. Fig. 7 shows the fractal dimension decrease from 1.25 to 1.24 as the surface is polished. It appears that the description of the progress of polishing and / or the profiles of the

structure of rough surface is useful way of describing structure of rugged surface quantitatively. The veins shown in Fig. 2 also have fractal structure.

7. Conclusion

It is a meaningful method describing polishing process by statistical analysis of many random grains.

The observation and analysis show that mechanical actions play an important role in this polishing process. The scratches which is much wider than grain size may be formed by the scraps of bonded grains. The swelled height of scratches has main influence on roughness.

The experiments shows that the oscillating speed, carrier speed, and their rate, and cylinder pressure have some influence on polishing results. The surface quality will be better when they are 120 1/min, 3 1/min, 14 and 70 kPa respectively in this research.

References

1. Mckeown, P.A., Cranfield, CUPE, 1987, The Role of Precision Engineering in Manufacturing of the Future Annals of the CIRP, 36/2:495-501
2. Yasui, H., Sato, K., 1994, Studies on High Flatness Polishing of Magnetic Disk Substrate with Al_2O_3 Abrasives, J. Japan. Soc. Prec. Eng., 60:128-132, 1631-1836
3. Mitsuhashi, M., 1990, Polishing of Substrate for Magnetic Memory Disk-Smoothing Characteristics of Substrate Surface Roughness, J. Japan. Soc. Prec. Eng., 56:311-316
4. Kawamura, Y., Kawakubo, Y., 1991, Surface Finishing of Coated Magnetic Disk Using Hydrodynamically Floating Abrasive Tape, J. Japan. Soc. Prec. Eng., 57:1579-1584
5. Tonshoff, H. K., Schmieden, W. V., etc., 1990, Abrasive Machining of Silicon, Annals of the CIRP, 39/2:621-635
6. Kasai, T., Karaki-Doy, T., 1990, Improvement of Conventional Polishing Conditions for Obtaining Super Smooth Surfaces of Glass and Metal Works. Annals of the CIRP, 39/1: 321-324

PHASE GRATING INTERFEROMETER GAUGE

D Mansfield & I K Buehring¹

Taylor Hobson Ltd
2 New Star Road
Leicester LE4 9JQ, UK

ABSTRACT

Taylor Hobson Ltd has recently launched the Phase Grating Interferometer¹ (PGI) gauge into the market. This metrological gauge, currently with a respective range and resolution of 10mm and 13nm, represents a profound change in gauging technology. The *modus operandi* of the PGI gauge is discussed together with a brief review of its performance.

INTRODUCTION

Briefly, the 'cornerstone' of the PGI gauge consists of a cylindrical holographic grating which is mounted on the end of a stylus arm such that its axis is coincident with the pivot axis. Directing a laser beam at normal incidence onto this grating then generates two 1st order diffraction beams; any movement of the stylus about the pivot is therefore transferred to a rotation of the grating. This in turn Doppler-frequency shifts the emerging diffracted 1st orders. Finally, super-imposing these diffracted orders generates an interference signal, the phase of which corresponds to the position of the stylus.

In this paper the principle of operation is discussed primarily from a physical optics vantage-point; this is followed by an overview of the gauge electronics. Finally, a review of the performance of the PGI gauge will be presented.

DIFFRACTION FROM A MOVING SINUSOIDAL SURFACE

It is useful to firstly consider the Fraunhofer diffraction resulting from a laser of Gaussian profile at normal incidence onto a sinusoidal surface. Referring to fig. 1, the near-surface field, $E(x,y)$, arising from such a planar incident wavefront is comprised of a Gaussian amplitude function with a sinusoidally modulated phase. It may be described by²:

$$E(x,y) = \frac{1}{r_o} \sqrt{\frac{2}{\pi}} \cdot \exp\left(-\frac{x^2+y^2}{r_o^2}\right) \cdot \exp\left(-i\frac{2\pi a_o}{\lambda}(1+\cos\psi)\cos\left(\frac{2\pi}{\lambda_s}(x-x_o)\right)\right)$$

The corresponding scalar Kirchhoff integral defining the diffracted far-field, E_ψ , may then be expressed:

$$E_\psi = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(x,y) \exp\left(i\frac{2\pi}{\lambda}x \sin\psi\right) dx dy$$

¹ Guidance Control Systems Ltd, Clockhill Wharfe, Atherstone Road, Hartshill, Warks CV10 0TB, UK

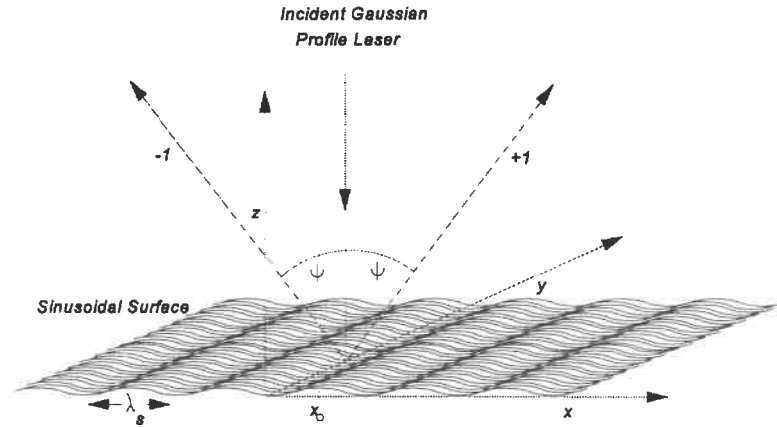


Figure 1 : Schematic of Sinusoidal Surface

As is well known, for such a periodic surface this integral collapses to zero for all values of ψ apart from those that satisfy the grating equation, $\sin\psi = q\lambda/\lambda_s$. For the q^{th} diffraction order, E_ψ may be expressed² as the convolution between the Gaussian point-spread function and the electric field E_q , where :

$$E_q = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp[-i\chi \cos(\xi - \phi)] \exp(iq\xi) d\xi$$

$$= i^q J_q(\chi) e^{iq\phi}$$

(where $\sin\psi = q\lambda/\lambda_s$, $\chi = 2\pi a_0(1 + \cos\psi)/\lambda$, $\xi = 2\pi x/\lambda_s$, $\phi = 2\pi x_0/\lambda_s$ and J_q is the integer Bessel function³ of order q).

Now, to examine the behaviour of this diffracted order when the grating has a lateral velocity v ($=dx/dt$), it is useful to expand the time dependency of the phase, Ψ of the q^{th} order diffracted field, E_q . Explicitly including the frequency term, $\exp(i 2\pi f_0 t)$, then :

$$\Psi = 2\pi(f_0 t + \frac{qx}{\lambda_s})$$

so that the instantaneous frequency is :

$$f = \frac{1}{2\pi} \frac{\partial \Psi}{\partial t} = f_0 + \frac{q\dot{x}}{\lambda_s} = f_0 + \frac{qv}{\lambda_s}$$

showing that such a movement of the sinusoidal surface results in a respective increase and decrease in the +1 and -1 order frequencies, in accordance with the well known Doppler relationship⁴. Finally, if the two 1st order beams are allowed to interfere, then the resulting signal is given by

$$I = |E_{+1} + E_{-1}|^2 \propto 1 + \cos \frac{4\pi}{\lambda_s} x$$

PGI GAUGE : INTERFEROMETER RÉSUMÉ

The underlying principles of the phase grating interferometer are described; whilst a rigorous description of the interferometer would clearly require the application of Jones matrices, these are avoided in the following description through the assumption of ideal components.

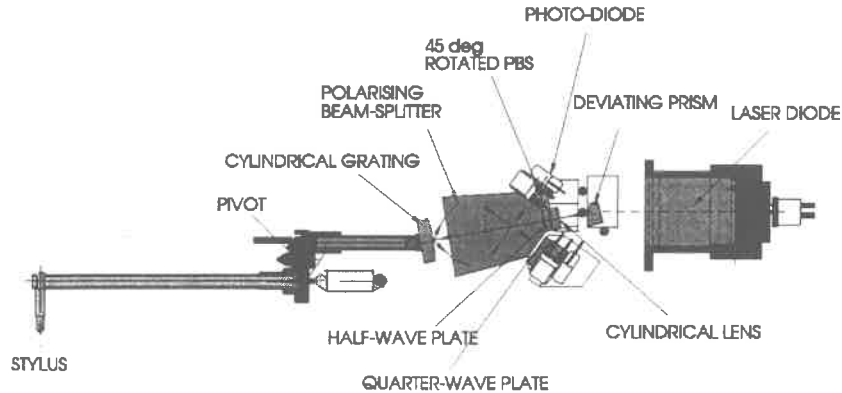


Figure 2 : Schematic of Phase Grating Interferometer Gauge

Referring to fig. 2, a laser diode of nominal wavelength (λ) 670nm is initially collimated. It has an elliptical cross-section with the long axis orientated out of the paper; additionally; it is plane polarised in the direction of the short axis. Passing through the deviating prism followed by the cylindrical lens it adopts a cylindrical wavefront (this pre-forming of the beam is necessary to ensure that the subsequently emerging 1st order diffracted wavefronts are quasi-planar). Prior to entering the main beamsplitter prism and then impacting on to the curved grating at normal incidence, the beam passes through a half-wave plate which rotates the plane of polarisation through 45°. The grating itself is a holographic cylindrical grating with a sinusoidal amplitude and a pitch, λ_g . The two emerging 1st order diffracted beams now have quasi-planar wavefronts and from above, have Doppler-shifted frequencies dependant upon the rotary movement of the grating about the pivot. The phases of these 1st orders are effectively independent of the laser wavelength, depending only upon the exact location (x) of the grating about the pivot and the grating pitch. Writing $\phi = 2\pi x / \lambda_g$ as above,

$$\underline{E}(\phi)^+ = \frac{E_o}{\sqrt{2}}(\hat{p} + \hat{s})e^{-i\phi}$$

$$\underline{E}(\phi)^- = \frac{E_o}{\sqrt{2}}(\hat{p} - \hat{s})e^{-i\phi}$$

Both 1st orders then enter the beamsplitter prism, undergo total internal reflection and are then mixed by the polarising beamsplitter to produce two orthogonally polarised beams :

$$\underline{E}_1(\Phi) = \frac{E_o}{\sqrt{2}} (\hat{p} e^{-i\Phi} + \hat{s} e^{i\Phi})$$

$$\underline{E}_2(\Phi) = \frac{E_o}{\sqrt{2}} (\hat{p} e^{+i\Phi} + \hat{s} e^{i\pi/2} e^{-i\Phi})$$

The term $e^{i\pi/2}$ allows for the relative phase advance of the s-plane electric field component when traversing the quarter-wave plate. Finally, the orthogonal fields in both exiting beams are mixed as a result of passing through the 45° -rotated polarising beamsplitters. These yield plane-polarised electric fields emerging onto the four detectors :

$$\underline{E}_1^s(\Phi) = E_o \hat{s} (e^{-i\Phi} + e^{+i\Phi})/\sqrt{2}$$

$$\underline{E}_1^p(\Phi) = E_o \hat{p} (e^{-i\Phi} - e^{+i\Phi})/\sqrt{2}$$

$$\underline{E}_2^s(\Phi) = E_o \hat{s} (ie^{-i\Phi} - e^{+i\Phi})/\sqrt{2}$$

$$\underline{E}_2^p(\Phi) = E_o \hat{p} (ie^{-i\Phi} + e^{+i\Phi})/\sqrt{2}$$

The signals from the corresponding detectors D_1^s , D_1^p , D_2^s and D_2^p are proportional to the incident intensities (given by the square of the modulus of the incident electric fields) which in turn finally yields the quadrature outputs :

$$D_1^s(\Phi) = \frac{I_o}{4} (1 + \cos\Phi) \qquad D_2^s(\Phi) = \frac{I_o}{4} (1 - \sin\Phi)$$

$$D_1^p(\Phi) = \frac{I_o}{4} (1 - \cos\Phi) \qquad D_2^p(\Phi) = \frac{I_o}{4} (1 + \sin\Phi)$$

Subtraction of the respective pairs of these then removes the dc terms so yielding $\sin\Phi$ and $\cos\Phi$ signals.

PGI GAUGE : FRINGE COUNT & INTERPOLATION

Referring to fig.3 and considering initially the fringe count, the 'Level Detection & Counter Logic' squares and latches the incoming sine and cosine signals on the positive going edge of a 10MHz clock signal. Comparators are then arranged to detect the zero-crossings, and additional logic based upon these squared signals then distinguishes (i) whether these zero-crossings are at 0° or 180° and (ii) whether the 0° zero-crossings correspond to an up-count or a down-count. The output is an 18bit fringe counter.

The tracking interpolator takes advantage of the identity

$$\sin\Phi\cos\theta - \cos\Phi\sin\theta = \sin(\Phi-\theta) \approx \Phi-\theta$$

to achieve 6 bit phase interpolated tracking. Briefly, the incoming $\sin\Phi$ and $\cos\Phi$ voltage signals are modified by the multiplying DACs to $\sin\Phi.\cos\theta$ and $\cos\Phi.\sin\theta$, the trigonometric multipliers being defined by the

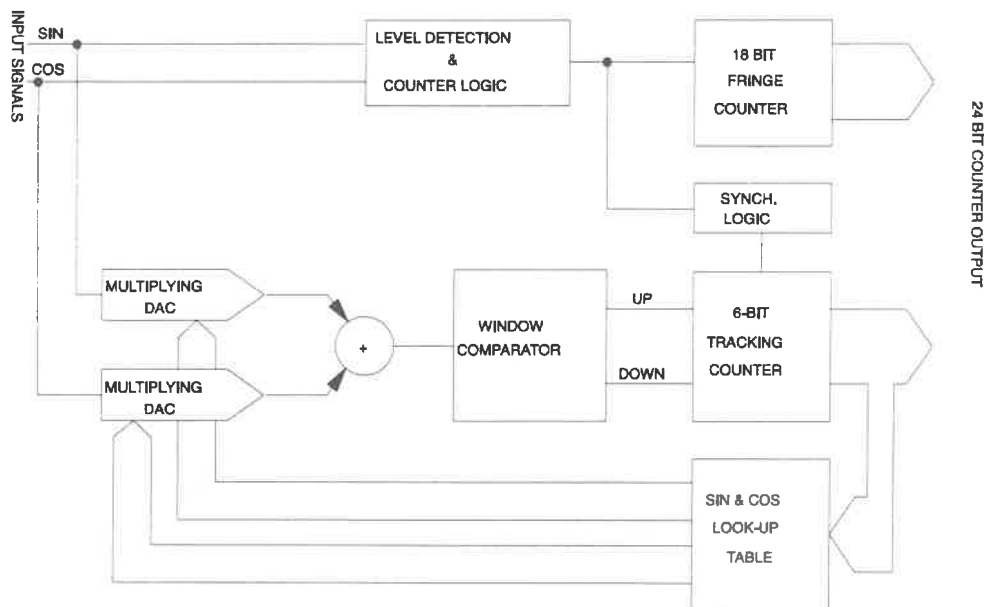


Figure 3 : Schematic of Fringe Counting & Tracking Electronics

LUTs. The voltage difference between the two terms is then input to the 'window comparator'; this increments, decrements, or leaves unchanged the current interpolated tracking phase, θ , depending on a comparison of $\phi - \theta$ and $2\pi/2^6$. These tracking logic signals are passed to the 6-bit tracking counter which both updates the interpolation angle for the LUTs and provides a 6 bit tracking phase interpolation output. Additional synchronisation logic ensures that the tracking interpolator's transitions $0 \sim 2^6$ (as the input phase, ϕ , passes through zero) are synchronised with the fringe counter incrementing or decrementing.

PGI GAUGE : PERFORMANCE

The PGI gauge may be fitted with a variety of different styli⁵; these allow both different ranges and components with differing restrictions to be measured. The standard 60mm stylus arm⁵ is assumed throughout the following. From above, the net output is shown to consist of 24bits. The 6 bit phase interpolation means that 1 LSB corresponds to a grating movement of $(\lambda_g/2)/2^6$, or 6.5nm corresponding to a stylus movement of 13nm. The current range of stylus movement is 10mm, corresponding to 20 bits. The bandwidth of the tracking interpolator allows tracking to be maintained at stylus vertical speeds in excess of 5mm/second - this ensures that tracking is maintained if the stylus 'drops' into a hole. Typically, over a 9mm stylus range, a gauge non-linearity of 0.1 μ m is achieved; this corresponds to 1.1×10^{-5} .

ACKNOWLEDGEMENTS

The considerable multi-disciplinary efforts required to develop the PGI into a fully engineered commercial product are acknowledged, in particular the mechanical and electronic design work of D Nettleton and D Onyon.

CONCLUSION

Taylor Hobson Ltd have recently introduced the PGI gauge. Based around a holographic cylindrical grating, it has been described with a particular emphasis on the relevant electro-optics. This compact metrological gauge, with its high bandwidth, range/resolution of 0.77×10^6 and a typical non-linearity of 1.1×10^{-5} , represents a profound change in gauging technology.

REFERENCES

1. Buehring I & Mansfield D, European Patent EP 0586454B1, 'Positional Measurement', priority date 30-05-91.
2. Mansfield D, 'Surface Characterisation via Optical Diffraction', Commercial Applications of Precision Manufacturing at the Sub-Micron Level', SPIE vol. 1573, Nov 1991.
3. Abramowitz M & Stegun A S, 'Handbook of Mathematical Functions', Dover, 1972
4. Drain L E, 'The Laser Doppler Technique', Wiley, 1980
5. Taylor Hobson Ltd, 'Form Talysurf Series II : 120mm PGI System', 1996

Wide-field scanning white light interferometry of rough surfaces

John Roth and Peter de Groot

zygo, Laurel Brook Road, Middlefield, CT 06455

Surface profiling with white light

Scanning white-light interferometry (SWLI) is a powerful tool for both surface form and roughness analysis. The technique shares the advantages of accuracy and high resolution with phase-shifting interferometry, but enjoys additional advantages by virtue of the broad emission spectrum of the source light. The broad spectrum makes it possible to profile surfaces with substantially greater roughness and surface departure than is practical with monochromatic interferometers.

A SWLI system involves an optical system similar to the one appearing in Figure 1. A computer accepts intensity data from an array of camera pixels while mechanically scanning the optical path difference (OPD) of an interferometer. The computer then processes the intensity data to determine the scan position for zero OPD at each pixel and hence the surface profile. The data processing involves either coherence envelope detection^{1,2} or spatial frequency analysis.^{3,4}

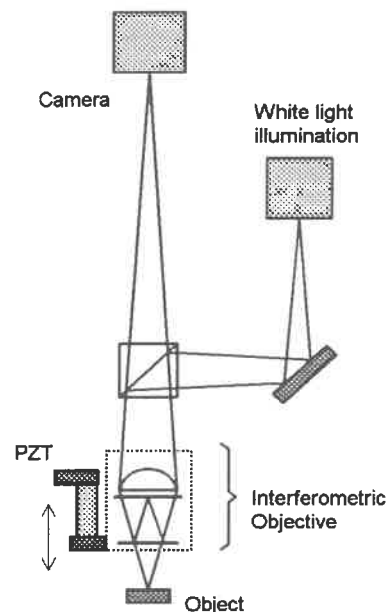


Figure 1: The scanning white-light interferometer employing a Mirau objective.

SWLI microscopy

All commercial implementations of SWLI are microscopes having a small field of view. Although larger fields of view are of great practical interest, there are fundamental reasons why the technique is best adapted to high magnifications (>2.5X). As the field of view increases, the rough-surface interference contrast decreases, eventually defeating the basic SWLI mechanism.

The problem may be understood qualitatively as follows. The coherence length L of an optical source can be thought of as the maximum OPD of an interferometer over which it is possible to still obtain interference. This length is

$$L = \lambda^2 / \Delta\lambda ,$$

where λ is the source wavelength and $\Delta\lambda$ is the spectral bandwidth. For example, Figure 2

shows an interference pattern for a 600-nm wavelength and a 125-nm bandwidth. A problem can arise when the surface roughness is comparable to the coherence length L of the light. The problem manifests itself as lost data or dropouts in the final image, because of inadequate signal strength.

At higher magnifications, the surface roughness is effectively resolved by the instrument, meaning that the range of topographical surfaces heights viewed by any one pixel is small. Thus the effective surface roughness for a single pixel is also small, and coherence length is less of an issue. As the magnification decreases, each individual pixel begins to see a larger range of surface heights, possibly covering a range comparable to the coherence length. This is why SWLI has been most successfully applied to microscopes.

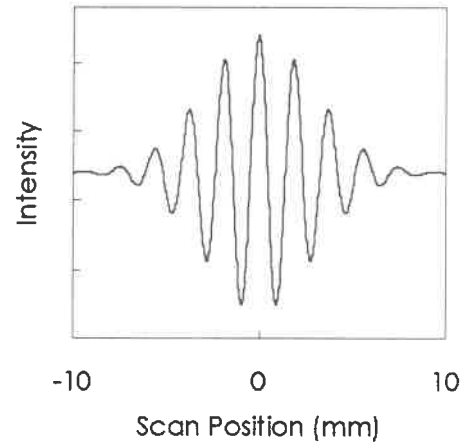


Figure 2: White light interferogram showing vanishing fringe modulation outside the coherence length.

Field enhancement with narrow-band filter

Given that the restriction on microscope FOV is closely related to the coherence length L , a reasonable strategy for reducing data dropout with rough surfaces is to reduce the spectral bandwidth of the source. An optical band-pass filter is a convenient way of adjusting the coherence length for this purpose. Figure 3, for example, shows a low-magnification SWLI image of paper using an 80-nm filter. The same object viewed with 125-nm light has 10% greater dropout.

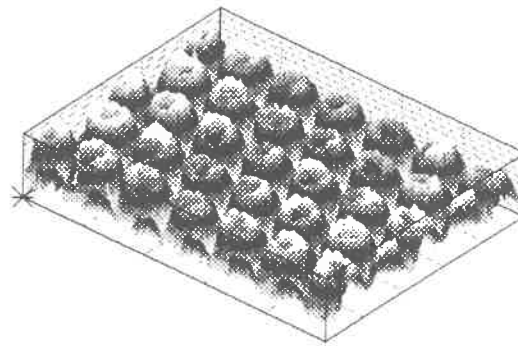


Figure 3: Narrow-band SWLI image of paper using a 2.5X Michelson objective. Because of the high surface roughness, broad-band SWLI typically requires a higher magnification and has significant data loss with this 2.5 X 1.9mm FOV.

Studies have also confirmed the advantages of narrow-band filtering for surface roughness analysis. With narrow-band light, the reported roughness more closely corresponds to stylus measurements over large surface areas. Table 1 summarizes these findings.

Surface type	True Roughness Ra $\pm$ 10%	125 nm SWLI result	80 nm SWLI result	Increase in valid data points
Turned	1.6 μ m	1.71 μ m	1.72 μ m	0 %
Milled	3.2	3.4	3.2	0 %
Turned	6.3	6.0	6.4	8 %
Milled	12.5	8.5	10.5	30 %

Table 1: Comparison of surface roughness measurements with two source filters.

Large fields with high resolution

Although spectral filtering is an effective way of improving SWLI measurements at low magnification, stitching together multiple measurements of adjacent areas on a surface results in even larger fields of view with no loss in resolution. This *field-stitching* technique involves moving the part under fixed optics through a series of small FOV measurements. Connectivity analysis combines the data from the separate measurements to produce a larger, composite three-dimensional topographic map.

In Practice, a computer aligns and positions the part automatically. At each measurement site, the x - y stage accurately positions the part, including a 10 to 20% overlap between sites. The computer records the height data together with the x , y , z information. The result an integrated image obtained with high resolution, large FOV and throughput



Figure 4: High lateral resolution SWLI image of a US quarter dollar.

consistent with automated inspection. The rapidly growing list of stitching applications includes lapping bars, fuel injector sealing surfaces, electronic packages, sensors, and medical devices.

References and notes

¹ N. Balasubramanian, U.S. Patent #4,340,306 (1982).

² P. J. Caber, "Interferometric profiler for rough surfaces," Appl. Opt. 32(19), 3438-3441 (1993).

³ P. de Groot and Leslie Deck, J. Mod. Opt. 42(2), 389 (1995).

⁴ US Patent Nos. 5,398,113 and 5,402,234.

Expanded Area Interferometric Surface Profiling with High Spatial Resolution

Paul J. Caber, James C. Wyant, and Joanna Schmit
WYKO Corporation, 2650 E. Elvira Road
Tucson, Arizona 85706, USA

INTRODUCTION

Obtaining large area measurements using optical interferometric metrology while retaining high spatial resolution has traditionally been very difficult to accomplish. While conventional interferometric systems using large aperture optics can indeed measure large areas, they typically provide very low spatial resolution. Interferometric microscopes can produce very high spatial resolution measurements, but are typically limited to measuring surface areas of very limited fields of view. It would be desirable from a metrology standpoint, to produce interferometric three-dimensional measurements of larger areas to allow the analysis of longer spatial wavelengths, while at the same time preserving the shorter spatial wavelength information. Also, it is often desirable to produce measurements of surface shape or form of larger samples that cannot be adequately profiled using conventional optical techniques. A practical method of accomplishing this has been developed that involves the stitching together of multiple high spatial resolution measurements of adjacent areas of a measured surface.

INTERFEROMETRIC TECHNIQUES

Two modes of operation are generally available for the interferometric surface profilers. For smooth surfaces the phase-shifting integrating bucket technique is generally used since it gives sub-nanometer height resolution capability. For rougher surfaces, a white light vertical scanning coherence sensing technique can be used to give nanometer height resolution over several hundred microns of surface height.

In order to obtain sufficiently high spatial sampling the field of view (FOV) of the measurement is typically very small. As an example, for a commonly used 740 x 480 element detector array and a 1.5X magnification between the sample and the detector array, the spatial sampling is 6.6 μm and the FOV is 4.9 x 3.2 mm. For a 50X magnification, the spatial sampling is 0.2 μm and the FOV is 0.15 x 0.10 mm. One approach for increasing the FOV, while keeping the spatial sampling constant, is to use a larger detector array, such as a 1K by 1K detector. Another approach, which can give a much larger FOV, is to stitch together several smaller FOV measurements to obtain a larger FOV high spatial resolution image.

THE STITCHING PROCESS

The stitching process involves fitting the measurements together in a global sense, or by matching the surface height and tilt over the overlap region. Care must be taken to make sure the measurements are precisely overlapped to minimize errors. The larger the overlap the easier it is to match data sets, but of course more data sets are required to get a given field of view. It has been shown that a 20 percent overlap gives a good trade off between having good repeatability and obtaining a large FOV with a minimum number of data sets. While the stitching approach is not new, the availability of precision stages with high-accuracy optical encoders, and fast microcomputers containing large amounts of memory makes the stitching approach much more attractive at the present time than it was a few years ago. However, there are still many sources of error that must be considered before one can obtain high resolution, large FOV images. Typical error sources include systematic or random single measurement errors, errors due to misalignment or mis-registration of the measurement system, and errors due to the stitching process itself. In this paper we will present a classification of important error sources and an analysis of these errors based on real data.

CLASSIFICATION OF ERRORS

Errors in the measurement process for a single measurement can lead to pronounced errors when these single measurements are stitched together for large FOV measurements. Typical measurement error sources are mis-calibration of the PZT or translating mechanism, leading to non-linearities or inaccuracies in the height data, magnification calibration errors, causing lateral dimensional inaccuracies leading to incorrect stage movement and subsequent mis-registration of the stitched images, and optical aberrations such as lens distortions and reference mirror curvature leading to improper stitching alignment or adjustment of tilt or height in the stitched overlap regions. Other single measurement errors that can affect stitched measurements include errors due to external environmental effects such as vibration and electronic interference, and improper setup of the measurement by the operator.

In stitching measurements it is important to precisely know the distance the sample is moved between measurements. Optical encoders make it possible to know the sample motion to better than one pixel. It is also important to ensure that the movement of the stage stays parallel to the rows and columns in the pixel array of the CCD camera. Precise adjustment of the rotation of the camera with respect to the stage is critical to producing good alignment in the stitched overlap regions. Along the same lines, parallelism and straightness of stage motion is also very critical, requiring stages with high-precision bearings and extremely tight tolerances on machined surfaces.

After completing a full set of measurements, the results are stitched together to form a large surface area topography. Several error sources arise when these individual sets of data are pieced together. Specifically, we concern ourselves with the influence of the

overlap of the measured areas, the number of measurements involved, and the order in which the sets of data are stitched together.

The measured areas need to include some area of overlap in order to adjust the relative tilt and height of each measurement. Three varying percentages of overlap were sampled to determine the optimal size of the shared regions. Greater overlap allows for more accurate fitting, whereas minimal overlap enables fewer measurements to be taken. In addition, as the number of measurements increases so does the error when the data is stitched together.

The influence of both the size of the overlap region and the number of measurements is shown in Table 1. Using the phase shifting technique, subtracting the effects of the reference wavefront, to eliminate systematic errors, and using autofocus prior to each measurement to minimize defocus errors, up to a 16 scan stitched measurement (4 x 4) of a roughness standard was taken at three varying percentages of overlap, 5, 20 and 35 percent. To analyze the repeatability of the stitched measurement the rms (root mean square) of the difference between the stitched measurements was calculated. From the rms values we see that the error for 5 percent of overlap is significantly larger than that for 20 and 35 percent. In order to examine the effect that the number of stitched measurements has on measurement error, we stitched together 3 x 3 and 2 x 2 measurement arrays from the existing data. For each percentage of overlap regions, the smallest error was achieved for the smallest number of stitched data, and for the largest overlap region the error was the smallest. From the table it is clear that as the overlap region decreases or the number of measurements increases the stitched measurement error gets larger. The 20% overlap region seems to be a good compromise between accuracy and speed.

Table 1: Average values of rms (nm) of difference between two stitched measurements.

Number of Files	5% overlap	20% overlap	35% overlap
Roughness standard RM S=0.5nm			
16 (4x4 files)	2.69	0.62	0.46
9 (3x3)	1.03	0.44	0.25
4 (2x2)	0.96	0.36	0.26
Roughness standard RM S=1.5nm			
9 (3x3)	1.55	0.54	0.37
4 (2x2)	0.96	0.51	0.34

Finally, the order in which the single phase maps are stitched together can make a difference, sometimes introducing a large error if care is not taken in designing the software, and sample roughness and structure can also influence the quality of the stitched measurement.

OTHER CONSIDERATIONS

Although the basic concept for stitching multiple area scans together is not new, and in fact has been demonstrated in one form or another for a number of years, the equipment necessary for successful and practical implementation has only recently become readily available. High precision stages are an absolute necessity, and only stages with precisely controlled motion, typically the result of high-quality crossed-roller bearings and precision machined surfaces, should be used. Accuracy of motion can be accomplished by using quality stages with high-precision lead screws and linear encoders.

Large area high-resolution cameras can not only greatly increase the FOV of a single measurement, but can also decrease the total number of scans, and thus improve the accuracy of a stitched measurement to achieve a larger FOV. Since excessive camera noise can cause errors that tend to propagate in a stitched measurement, low noise digital cameras are also helpful in providing good results.

Since the size of large FOV stitched measurements can generate large amounts of data, computers with large amounts of RAM become necessary for practical handling of these large arrays. Computers with up to 1/2 G byte of memory have been used to process large stitched measurements. The overall computational speed of the computer system is also a limiting factor on large area measurements, and only recently have computers been fast enough to make some of these measurements practical.

The ability to stitch surfaces to obtain high lateral resolution over a large field of view should further increase the applications of phase-shifting and vertical scanning interferometric profilers.

REFERENCES

1. J.C. Wyant and J. Schmit, "Large field of view, high spatial resolution surface measurements," Proc. of the 7th Int. Conf. On Metrology and Properties of Engineering Surfaces, pp. 294-301, (1997).
2. K. Creath, "Phase-shifting interferometry techniques," in Progress in Optics XXVI, E. Wolf, ed. (Elsevier Science, 1988), pp. 357-373.
3. P. J. Caber, "An interferometric profiler for rough surfaces," Appl. Opt. 32(19): 3438-3441, (July 1993).
4. E. R. Cochran and K. Kreath, "A method for extending the measurement range of a two-dimensional surface profiling instrument," Proc. SPIE, (August 1987).

RAPID DATA ACQUISITION AND ANALYSIS FOR PART PROFILING

Richard Jennings
ICAMP, Inc.
Bolton, CT

While Profile is a measurement most often associated with localized surface variation, a related measurement in the Dimensioning & Tolerancing Standard, Y14.5M-1994, is generating considerable interest in its ability to “control form or *combinations* of size, form, orientation, and location” of features and entire parts. We refer to Section 6.5 on Profile Control and Profile Tolerancing.

While a variety of inspection devices, such as computer controlled Coordinate Measuring machines (CMMs) are now capable of very rapid collection of 2D and 3D *point* data, these devices often incorporate proprietary measurement software which operates on geometric approximations of the part being inspected.

These approximations are known as “Substitute Geometry” and are created by the measurement software as the perfect form which most closely approximates the actual part or feature. The measurement software then compares the Substitute Geometry to the engineering design and reports any variations in size, shape, and position between Substitute Geometry and the engineering design. It should be obvious that comparing two forms of known size and shape is easier than comparing a “cloud of data” to a geometric form.

If the part being inspected is perfect, Substitute Geometry will provide perfect results. Unfortunately, few parts are perfect and inaccuracies can be introduced in measurement computations. So long as design tolerances were “loose”, such inaccuracies absorbed only a small part of the tolerance “budget” and were not noticeable.

However, dimensional tolerances for parts as diverse as bearings, machine tool inserts, and computer disk components are “shrinking” rapidly; as an example, the National Institute of Standards and Technology (NIST) has reported that tolerances for castings have shrunk by a factor of four in the last few years. As tolerance ranges shrink, computation techniques that were once acceptable may now yield misleading results.

This paper describes the development, testing, and integration of an independent software measurement, analysis, and visualization system, incorporating Profile Tolerancing, which can be linked to any non-contact, scanning, or contact probe sensor which is capable of electronically recording and transferring discrete point values.

Rather than utilizing Substitute Geometry, this methodology first creates a mathematically exact solid model corresponding to the engineering design. This model can be used for nominal comparisons. Solid models for analysis are particularly attractive since Y14.5M is based on material condition and specifies analysis and computations which report on material values.

When tolerances are used, this methodology creates solid "shells" parallel at all points on the shell surfaces to the engineering nominal model but offset from the nominal surfaces at distances equal to the tolerance values.

When development started, a primary goal was to avoid the "point-to-point" comparison methods of most inspection systems which can yield serious errors in reported values if the part is not perfect. If the *complete set* of measured points could be compared to a solid model, more accurate measurements would be possible.

An example of the greater accuracy provided is in the elimination of Cosine Error; this type of error is introduced in contact probe CMMs when the probe does not approach the part surface along a path normal to the part surface at the point of contact. This error can be significant on a surface with rapidly changing slope. While most CMMs are performing "point-to-point" measurements, the use of a solid model allows the software to "simultaneously" compute the magnitude of a vector from each point in the data set (rigid body) to all points on the model surface and to float the data set (in 3-space) over the model and to find the position for which the sum of the squares of the distance of each point from the model surface ($d_1^2 + d_2^2 + d_3^2 + \dots$) is minimized. This procedure is called a Least Squares Best Fit.

By mathematically "attaching" each measured point to the closest surface of the solid model, the surface normal can be computed. The vector magnitude of the surface normal is the deviation for that point. Since the deviations are from the surface of a solid model, the signed value represents excess (+) or deficient (-) material. Deviations whose vector normal magnitude penetrate the tolerance shells can be specially flagged for further analysis and observation.

By establishing parallel tolerance shells, the software system can ask, and answer, a fundamentally new set of questions. Rather than computing feature based measurements (flatness, diameter of a cylinder, etc.), the software asks, "Can the point data set representing the inspected part be translated and rotated, as a rigid body, into a position in which all points fit inside the tolerance shells?" By definition, if the point set can be fitted between the tolerance boundaries, the part is acceptable.

Process engineers can also answer questions such as: "How much smaller could the tolerance shells be and still contain the part?", or "How much larger would one or more tolerance shells need to be to pass the part?"

This methodology subsumes several measurement methods defined in Y14.5M-1994 including Maximum Inscribed and Minimum Circumscribed circles and cylinders for hole patterns and bosses and provides measurements consistent with GD&T criteria.

The next step was to provide simple, easy to understand (for non-inspection personnel) visual representation of the point deviations. It was straightforward to represent each measured point as a deviation "whisker" with one color representing points higher (larger) than nominal and another representing points lower (smaller) than nominal. To make it easier to see patterns in the deviations, whisker lengths are scaled independently of the model dimensions.

While this implementation proved useful for design and process engineers, it was almost too flexible for long time inspection practitioners attuned to collecting inspection data on parts constrained in "datum" positions. To satisfy such concerns, it was a straightforward step to simply superimpose the tolerance "shells" electronically over the part data, without allowing any movement, and check for points penetrating the shells. This procedure, of course, assumes that the part was perfectly positioned in the inspection space.

This "fixed" procedure provides a very accurate representation of part condition **as-inspected**; but the as-inspected position and orientation may contain errors. To "squeeze" out such alignment problems, the same Best Fit algorithms described earlier can be introduced to allow the model to translate and rotate in X,Y, and Z. An iterative process computes the sum of the least squares of all of the deviations and converges on a minimum score. The position yielding the lowest score is considered to be the Best Fit Position for the part and this position can be used for all analysis functions.

To assist in process evaluation, or to Best Fit against datums, the user can "freeze" any direction of motion or can choose to fit against any, some, or all surfaces.

In this methodology, features in the solid model, rather than inspection fixtures, are used as Datum Feature Simulators. Since the model (theoretical) features are perfect, there is no need to allocate the "Gagemaker's Tolerances" usually recommended for hard datums; this may free up an additional 5-10% of the total tolerance for use in the part itself. Experiments have shown that allowing the part to "datum against itself" can increase part yield in inspection.

Datums are an inspection artifact and, in some cases, can be useful in assembly. But to accept all good parts while rejecting all bad parts, this methodology is best used to literally "fit the part against itself". By allowing the measured part to float between tolerance shells without regard to datums, the analysis will find all parts that can fit between the tolerance boundaries.

This mathematical system may pass parts which would fail the hard gage test if the data set is sparse; but with the advent of high speed scanning probes, video/vision systems, and laser systems, the amount of data which can be collected provides a high level of confidence in the computations which are generated.

The down side of large data sets is, of course, the possible protracted run times required to process the data in such compute intensive activities. Even with high speed personal computers, great care must be taken in coding analysis algorithms to minimize run time while maintaining accuracy.

Some surprising secondary observations have evolved from the development of solids based part profiling systems. The first was that reliance on Part Profiling would allow the design engineer to reduce the quantity and detail of design tolerancing needed to define the part while actually *increasing* control over part variation. The second was that inspection programming effort could be significantly reduced since only a "cloud" of data points is needed to drive a solids based analysis. The third was that analysis could be physically separated from the inspection process since the data set was preserved and could be analyzed at any time; separating analysis from inspection would also mean that more parts could be inspected with the same inspection equipment. The fourth was that basic inspection measurements such as True Position at Maximum Material Condition (MMC) were easily added to solids based systems.

The major force holding back acceptance of these systems is the reluctance of many inspection vendors to open their systems to third party software. The data needed for solids based analysis is the same point data gathered by all inspection machines and used in their feature measurements; but currently most inspection systems throw those points away after computing feature measurements such as Diameter.

This Profile Tolerancing Analysis software has been successfully linked to, and used in production with, commercial video/vision systems from Micro-Vu, OGP, RAM, and View; contact probe systems from Carl Zeiss, Brown & Sharpe, Leitz, Mitutoyo, and Sheffield; electrostatic probes from EMD and Extrude Hone; and profile machines such as TSK.

Both commercial and academic organizations are developing "tolerance checking" systems and the advantages of such systems would seem to indicate that they will replace Substitute Geometry based systems in the future for dimensional inspection.

CALIBRATION PROBLEMS IN SURFACE METROLOGY

T. V. Vorburger, R. G. Dixon, V. Tsai*, T. Doi⁺, J. F. Song, T. B. Renegar, and J. Fu
National Institute of Standards and Technology
Gaithersburg, MD 20899 USA

* University of Maryland, College Park, MD 20742 USA

⁺ Permanent Address: National Research Laboratory of Metrology
1-1-4 Umezono, Tsukuba, Ibaraki 305 JAPAN

ABSTRACT

Calibrations of surface finish measurements in our laboratory rely on a set of master step-height specimens ranging in size from 29 nm to 152 μm . These step-height specimens are calibrated in turn by different techniques depending on the step height itself. In order to determine precise values for step heights smaller than 1000 nm, we have established a goal to calibrate several new step-height specimens in this range by three different independent techniques. These are phase shift interferometric microscopy, a calibrated atomic force microscope (C-AFM), and a high-resolution stylus instrument calibrated by another calibration step. If we can measure precise values of the step height by these different techniques and achieve good agreement, the step-height specimens would then become precise calibration masters.

Each of these instruments has different advantages. The stylus technique is an excellent step-height comparator. The vertical displacement transducer of our instrument is a highly linear and stable LVDT (linear variable differential transformer) with subnanometer noise resolution, but this transducer must be calibrated in terms of a wavelength of light. Furthermore, as it is a mechanically contacting technique, there is a potential for surface damage. The other two instruments closely tie the step-height measurements to the wavelength of light. The phase measuring interferometer uses a white light source with a narrow band filter and provides a completely non-contacting measurement of the step height, whose value is calibrated in terms of the mean wavelength of the filter. The C-AFM combines a contacting or proximity probe with a calibrated transducer output. The vertical motion of the specimen is driven by a piezoelectric stack and measured by a capacitance gauge integrated into the motion package. The output of the capacitance gauge is calibrated routinely by a HeNe laser interferometer.

The first sets of measurements have been performed on a set of quartz steps with nominal heights of 15 nm, 90 nm, and 1000 nm. To this point, the most repeatable results have been obtained with the stylus instrument. The measurements with the phase measuring interferometric microscope show variations due to several causes, in principal ones being defocus of the specimen and tilt of the specimen with respect to the optic axis. The calibration of the vertical travel of the C-AFM is precise. Nevertheless, significant variations in the measured step-height values occur. These may be arising from out-of-plane motion of the specimen, variations due to changes in the lab environment, or tip-surface interactions that cause changes in the characteristics of the cantilever force sensor.

Measurements of Pitch, Height, and Width Artifacts with the NIST Calibrated Atomic Force Microscope

R. Dixon¹, V. W. Tsai², T. Vorburger¹, E.D. Williams³, X.S. Wang³, J. Fu¹, R. Köning¹

¹ National Institute of Standards and Technology, Gaithersburg, MD 20899

² Dept. of Material Engineering, University of Maryland, College Park, MD 20742

³ Dept. of Physics, University of Maryland, College Park, MD 20742

Introduction

Use of the atomic force microscope (AFM) in industrial metrology is rapidly increasing. Properties commonly measured in such applications are feature spacing (pitch), feature height (or depth), feature width (critical dimension), and surface roughness. To achieve high accuracy, the scales of an AFM must be calibrated. The use of a standard is normally the most appropriate means of doing this. Presently available standards for this purpose are calibrated using stylus instruments and optical techniques. The effectiveness of this approach, however, is limited by the differences in the working ranges of the various techniques and by questions of methods divergence. As the repeatability of commercial instruments continues to improve, the importance of accurate calibration standards will increase.

We have designed and are developing the calibrated AFM (C-AFM) to calibrate standards using an AFM. As described elsewhere [1-3], this instrument has metrology traceable to the wavelength of light for all three axes. To accomplish this, a flexure translation stage, heterodyne laser interferometers, and a digital-signal-processor based closed-loop feedback system are used to control the x-y scan motion. The z-axis translation is accomplished using a piezoelectric actuator with an integrated capacitance sensor, which is calibrated using a heterodyne laser interferometer. When fully developed, this instrument will be a calibration tool for scanning probe microscope standards. Specifically, our primary calibrations are expected to be of combined pitch/height, or three dimensional magnification standards. In the present work, we describe the current status of the C-AFM and the use of the instrument for measurements of pitch, height, and width. With respect to height measurements, we will place the most emphasis on some of our recent efforts using the C-AFM to study Si samples with single atomic steps.

System Goals and Status

Near term measurement goals for the C-AFM system include: (1) the capability of calibrating sub-micrometer pitch artifacts with an expanded uncertainty (2σ) of three nanometers; (2) the capability of calibrating sub-micrometer height artifacts with an expanded uncertainty of 0.3 nm plus 1% of the height; (3) the measurement of the top width of sub-micrometer vertical features with an expanded uncertainty of 20 nm [4].

The z-stage on the C-AFM system is being replaced, and the new stage will have significantly lower tilt through the range of travel. We are reevaluating the tilting motions of the x-y stage. When this is complete, we will be able to reduce our uncertainties arising from Abbe errors, which will be necessary to accomplish our goal of calibrating sub-micrometer pitch artifacts with an expanded uncertainty less than 3 nm. Our testing shows that the roll-pitch-yaw of the x-y stage to be less than one arc second over the range of travel of 50 μm . Prior to the current modifications, we completed a comparison of pitch, height, and width measurements made on the C-AFM with those obtained by other methods. When the modifications are complete, we will repeat this process.

We are also developing a sample positioning system in the C-AFM to facilitate the location

of features of interest within the scan range of the instrument. The short term solution involves an assembly which allows for mounting a micro-manipulator near the side of the metrology frame. This manipulator is used to position the sample on the specimen platform. A longer term solution under consideration involves the installation of a separate translation system for relative motion between the x-y stage and the metrology frame.

System Performance and Plans

Pitch Measurements

In the previous testing of the system, we estimated the standard uncertainty (1σ) for pitch measurements to be 7 nm for large pitch values (on the order of 10 μm) and approximately 3 nm for smaller, sub-micrometer pitch values. The uncertainty at larger values is thought to arise mainly from unwanted motion between the AFM tip and the metrology frame during the measurement. For smaller pitch values, the 3 nm uncertainty limit arises from the uncertainty in the Abbe error. As mentioned above, we expect to reduce this uncertainty contribution once we have completed additional testing of our x-y stage. Our current testing has shown that the maximum values of roll, pitch, and yaw of the stage, over the range of travel of 50 μm , are about one arc second. It also shows that we may be able to select a subset of the range, in which our sub-micrometer pitch calibrations will be carried out, over which the tilting is less than one quarter of one arc second. This means that the error arising from a one millimeter Abbe offset would be about one nanometer. Other sources of uncertainty are the potential for cantilever flexing or tip deformation during measurement and polarization mixing in the laser interferometers.

Height Measurements - Large Scale

The standard uncertainty of C-AFM step height measurements of steps ranging from 18 nm to 920 nm was estimated to be approximately 1.9%. For these measurements, the dominant uncertainty component arises from the variability of the results under different measurement conditions. We believe this variability is due partly to cantilever flexing and to cantilever-surface interactions. Out-of-plane motion (i.e. z-straightness of the x and y axes) of the x-y translation stage and changes in laboratory conditions (temperature and humidity) may also play a role. Smaller sources of uncertainty include variability and linearity of the capacitance gauge and Abbe offset.

Height Measurements - Small Scale

We have recently focussed much of our attention on the measurement of Si single atomic step artifacts. There is presently no height standard in the sub-nanometer range commercially available. The type of samples we have prepared for this work are promising candidates for AFM calibration at this scale. Silicon surfaces can be prepared in UHV with steps determined by the lattice spacing having expected heights from 0.14 nm to 0.31 nm, depending upon the crystal orientation [6,7]. We have prepared a variety of Si surfaces with single atomic steps using an ultra high vacuum (UHV) system constructed at the University of Maryland [8,9], and we have characterized these surfaces using UHV scanning tunneling microscopy (UHV-STM), commercial AFMs, and the C-AFM [10]. Measurements using the C-AFM are performed after the growth of the native oxide. As observed by the C-AFM, the basic step structure on these surfaces with native oxide is stable under ambient conditions for periods exceeding one year after initial sample preparation.

We believe that Si(111) surfaces having steps of about 0.31 nm are probably more appropriate as standards than the 0.14 nm steps. The latter may be too small relative to the noise resolution of some instruments. Therefore, our measurements using the C-AFM have been focussed on measuring the ~ 0.31 nm steps on Si(111) surfaces. We are using both the contact and

intermittent contact modes of operating the C-AFM in these investigations. The present discussion will focus on some data taken in the intermittent contact mode. Results from two sets of measurements are described here. The first set consists of twenty successive images of the same step without changing location, and the second set consists of images of twenty adjacent steps on the sample. Results from these two sets of data, with the step height calculated by three different algorithms as described below, are shown in table I.

The most appropriate method for determining a step height from the C-AFM data is the subject of ongoing investigation. Reduction of noise by averaging and removal of systematic errors are important considerations. We are presently exploring different approaches to the analysis, but for this work we have chosen to apply three methods to our data. Two considerations in analyzing the data are (1) the method of areal averaging and (2) the definition of the step height. The first two methods involve averaging by profiles with different definitions of the step height. The third method uses true areal averaging and is easily employed with alternative definitions of the step height. In method I, the step height is taken to be the difference in the average height of selected regions of the upper and lower surfaces. In method II, lines are fit to the upper and lower portions of the profile (excluding the step edge), and the step height is taken as the difference in these lines extrapolated to the edge of the step. In method III, planes are fit to the upper and lower portions of the step (excluding the edge) and these are extrapolated to a point in the middle of the step edge, although alternative extrapolation points could be chosen.

The expected value of the Si(111) step based on the lattice constant of the underlying bulk silicon is 0.314 nm. Other investigators have measured the height of the step features of Si(111) in UHV to be 0.31 nm $\pm$ 0.01 nm using LEED [6], and 0.34 nm $\pm$ 0.03 nm using x-ray diffraction [11]. More recently, Suzuki et al. [12] have measured the step heights of such samples in ambient with a native oxide present to be 0.35 nm $\pm$ 0.03 nm using commercially available AFMs and their own calibration standards. The average value obtained from all of our measurements of these steps, including those shown in table I, is about 0.29 nm. At present, the differences between these various results are not fully understood. The same step repeatability of our measurements, as evident from table I, is about 0.006 nm, and the non-uniformity, at least of adjacent features, is about the same. The observed level of agreement between our three methods of analysis makes it clear that, although the definition and calculation of step height are important considerations, the uncertainties associated with our analysis algorithms are not a factor in understanding the observed discrepancies in step height. We have also tested the algorithm of Suzuki, et al. [12] with our data and found little difference. However, our overall reproducibility, including values from different samples, using both piezoresistive and optical lever AFM heads, and both contact and intermittent contact modes is about 0.04 nm (1σ). We are working to improve the reproducibility by searching for unknown systematic errors in our measurements, and we plan to study the morphology of the surfaces as the native oxide forms.

Width Measurements

The accuracy of C-AFM width measurements is presently limited by our knowledge of tip geometry. We are exploring the use of tip characterizer samples to estimate the tip size and correct the apparent feature width for the increase due to the probe. By using nanometer-sized colloidal gold particles to estimate the width of our tips, we were able to perform top-width measurements with an estimated 15 nm standard uncertainty [5]. We plan to use other types of characterizers and tips to further improve the accuracy of C-AFM width measurements.

Currently, we are involved in a collaborative effort, involving both NIST and Sandia National Laboratories, to compare measurements of linewidth on the same sample obtained by various

techniques, including AFM, SEM, optical microscopy, and electrical resistance. The samples are prepared using silicon-on-insulator (SOI) substrates and preferential (KOH) etching to obtain lines with vertical sidewalls [13]. Figure 1 shows a C-AFM image of a line on one of these samples, along with its intersection with one of the electrical taps used for the resistance measurement. Although the sidewalls of these lines are actually vertical, the shape of the C-AFM tip causes an apparent slope to the sides of the line. Our measurement [14] of the linewidth, without complete tip characterization and correction, is about 100 nm larger than the electrical resistance value. We plan to perform more measurements and use tip characterizers to clarify the difference between the results for these samples. Additional evaluation of the uncertainties in the calculation of the linewidth from electrical resistance is also underway.

Conclusions

We have described the current status of the C-AFM and the performance of the system for measurements of pitch, height, and width. Additionally, we have discussed the ongoing modifications and testing, and our plans for the further development of the system. Our use of the C-AFM to study Si single atomic step samples as potential calibration standards has been discussed in some detail. Once the differences between our measurements and those of others are more fully understood, we believe that Si single atomic step samples will have considerable potential as 'self-calibrating' standards for AFM. Investigation of this potential height standard will continue in parallel with development of a combined pitch/height, or three axis magnification standard. We expect both of these types of standards to play an important role in the continued development of AFMs as metrology instruments.

Acknowledgments

This work was supported largely by the National Semiconductor Metrology Program (NSMP) and National Advanced Manufacturing Testbed (NAMT) at NIST. We thank C. J. Evans, E. C. Teague, R. S. Polvani, J. S. Villarrubia, and J. Dagata of NIST for many very helpful and informative discussions.

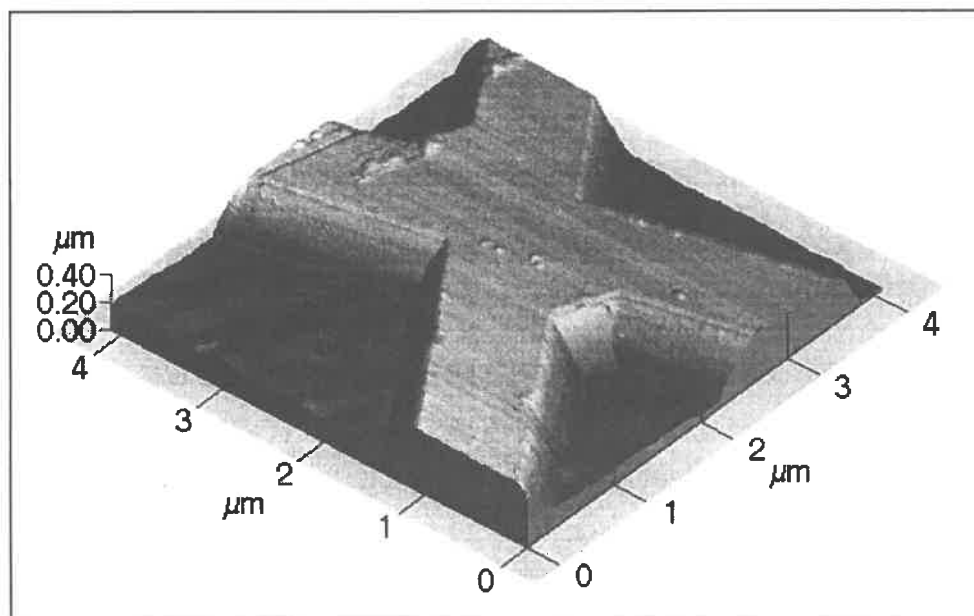


Figure 1. Topographic map, as measured by the NIST C-AFM, of a NIST/Sandia SIMOX linewidth sample showing an $\sim 0.8 \mu\text{m}$ specimen line intersecting an electrical tap. (Note: The specimen line runs from roughly the top of the figure to the bottom.)

Table I. Step height results from two data sets on a silicon (111) surface using the C-AFM, including standard uncertainties (1σ - Type A) derived from the variations in the measurements.

	Analysis Method 1	Analysis Method 2	Analysis Method 3
Data Set I (Same step images)	$0.272 \text{ nm} \pm 0.005 \text{ nm}$	$0.266 \text{ nm} \pm 0.006 \text{ nm}$	$0.266 \text{ nm} \pm 0.006 \text{ nm}$
Data Set II (adjacent step images)	$0.273 \text{ nm} \pm 0.009 \text{ nm}$	$0.269 \text{ nm} \pm 0.004 \text{ nm}$	$0.269 \text{ nm} \pm 0.004 \text{ nm}$

References:

- [1] R. Dixon, T. V. Vorburger, P. J. Sullivan, V. W. Tsai, T. H. McWaid, *ASPE Proceedings* Vol. 14, 375 (1996).
- [2] J. Schneir, T. H. McWaid, R. Dixon, V. W. Tsai, J. S. Villarrubia, E. D. Williams, E. Fu, *SPIE Proceedings* 2439, 401 (1995).
- [3] J. Schneir, T. H. McWaid, J. Alexander, and B. P. Wilfley, *J. Vac. Sci Technol. B*, 12 (6), 3561 (1994).
- [4] B. N. Taylor and C. E. Kuyatt, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Tech. Note 1297 (1994).
- [5] R. Dixon, J. Schneir, T. McWaid, N. Sullivan, V. W. Tsai, S. H. Zaidi, S. R. J. Brueck, *SPIE Proceedings* Vol. 2725, 589 (1996).
- [6] R. J. Phaneuf and E. D. Williams, *Phys. Rev. Lett.* 58, 2563 (1987).
- [7] E. D. Williams, R. J. Phaneuf, J. Wei, N. C. Bartelt, and T. L. Einstein, *Surf. Sci.*, 294, 219 (1993).
- [8] R. J. Phaneuf and E. D. Williams, *Phys. Rev. B* 35, 4155(1987)
- [9] Y. N. Yang, E. S. Fu and E. D Williams, *Surf. Sci.* 356, 101 (1996).
- [10] V. W. Tsai, T. V. Vorburger, P. J. Sullivan, R. Dixon, R. Silver, E. D. Williams, J. Schneir, *ASPE Proceedings* Vol. 14, 212 (1996).
- [11] D. Y. Noh, K. I. Blum, M. J. Ramstad, and R. J. Birgeneau, *Phys. Rev. B* 44, 10969 (1991).
- [12] M. Suzuki, et al., *J. Vacuum Sci. Technol. A* 14, 1228 (1996).
- [13] M. W. Creswell, J. J. Sniegowski, R. N. Ghoshtagore, R. A. Allen, W. F. Guthrie, A. W. Gurnell, L. W. Linholm, R. G. Dixon, and E. C. Teague, *Jpn. J. Appl. Phys.* 35, 6597 (1996).
- [14] J. S. Villarrubia and R. Dixon, unpublished results (1997).

ATOMIC FORCE MICROSCOPY FOR PROFILER STYLUS QUALITY CONTROL

Hans H. Gatzert, J. Chris Maetzig, and Heiko Siekmann
Institute for Microtechnology
Hanover University, Hanover, Germany

Abstract—In surface profiling, the diamond stylus radius and cone angle are important parameters which have to be tightly controlled. Until recently, for measuring the stylus geometry, only qualitative analysis methods were available. This paper presents an approach how to measure the stylus geometry, in particular the stylus sphere radius, using Atomic Force Microscopy (AFM). As it turns out, AFM is very well suited for quantitatively analyzing the topography for profiler diamond styli.

I. INTRODUCTION

The standard method for analyzing surface roughness is mechanical stylus profiler measurement. To guaranty the measurement of comparable roughness values, the tip geometry should be closely controlled. Furthermore, a trend towards finer stylus geometries is required to cope with decreasing roughness due to today's advances in ultraprecision surface machining. However, so far there has been no high resolution quality control tool for determining the stylus tip's microtopography. This need may be filled by means of Atomic Force Microscopy (AFM) allowing for the first time a quantitative analysis of the stylus tip's key dimensions.

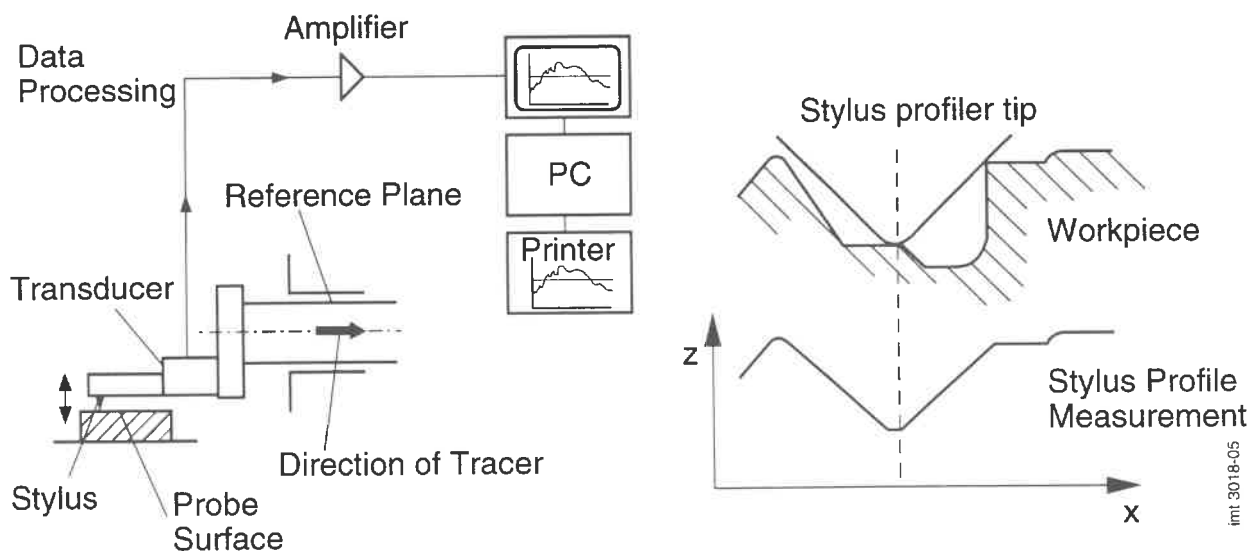


Fig. 1: Stylus profiler
a. Schematic (left)
b. Profile distortion caused by the stylus' geometry (right)

II. STYLUS PROFILER MEASUREMENT TECHNIQUE

Fig. 1a depicts the schematic of a stylus profiler. For executing the measurement, a stylus is lowered on the probe's surface until there is contact. A traverse unit traveling at constant velocity draws the stylus across the surface. While tracing the probe surface the traverse shaft provides a straight reference datum. A transducer picks up the stylus' vertical displacement, a computer system stores the amplified displacement data and allows the display of a two-dimensional (2-D) surface profile as well as the calculation of respective roughness values as there are R_a and R_z . Measurement parameters and stylus geometries like cone angle and tip radius are standardized [1, 2].

The styli are made of diamond, a standardized stylus is a cone of a 90° or 60° angle and a spherical tip with a 10, 5, or 2 μm radius respectively. The tolerances for the included cone angles are $\pm 5^\circ$, for the tip radii 30 to 50 percent of the nominal radius. To measure ultrasmooth surfaces, smaller radii in the sub- μm range are in use although not yet standardized. Furthermore, it is highly desirable to further reduce specifically the sphere radius tolerances. Fig. 1b demonstrates the importance of a small tip radius and cone angle as well as a tight control of the stylus' profile when measuring surfaces with short wavelength topography features. The finer the stylus, the smaller the wavelength at which cutoff starts to occur [3]. On the other hand, even if cutoff does occur, comparable measurement data are achieved, as long as the stylus geometry follows a standard. A difference in cone geometry will result in a different representation of the surface topography and comparing measurements become impossible.

Today, the means for a quality control on styli are rather limited. Typically, taking Scanning Electron Microscopy (SEM) micrographs of the stylus provides for a qualitative quality control, limiting the stylus analysis to aspects like defects and fractures, and, best case, to a quantitative analysis of the cone tip radius which by chance shows in profile. A systematic quantitative analysis of the sphere's various planes, which is most desirable, is not available.

III. AFM MEASUREMENT TECHNIQUE - CONTACT AND NON-CONTACT

During the recent years, AFM analysis has helped to tackle quite a view micromasurement tasks unresolved beforehand. It also looks like a very interesting approach to resolve the diamond stylus measurement challenges. To investigate which AFM measurement technique lends itself best for analyzing stylus geometry, both "contact" mode and "non-contact" mode AFM measurements are executed. Fig. 2 shows the schematic of the AFM system.

Fig. 2a depicts the contact mode measurement principle. A silicon nitride tip with a length of 4 μm and a tip radius smaller than 50 nm located at the end of a cantilever contacts the probe. The mirror finish cantilever back surface reflects a laser beam and directs it towards a four quadrant photo detector. A set of x- and y-direction piezo actuators scan the probe's surface. The repelling force invoked by the probe surface on the tip deflect the cantilever according to the local surface topography. A z-direction piezo adjusts the cantilever elevation, keeping the cantilever deflection constant at all times by controlling the z-height based on the photo detector output. Thus the z-direction piezo voltage signal

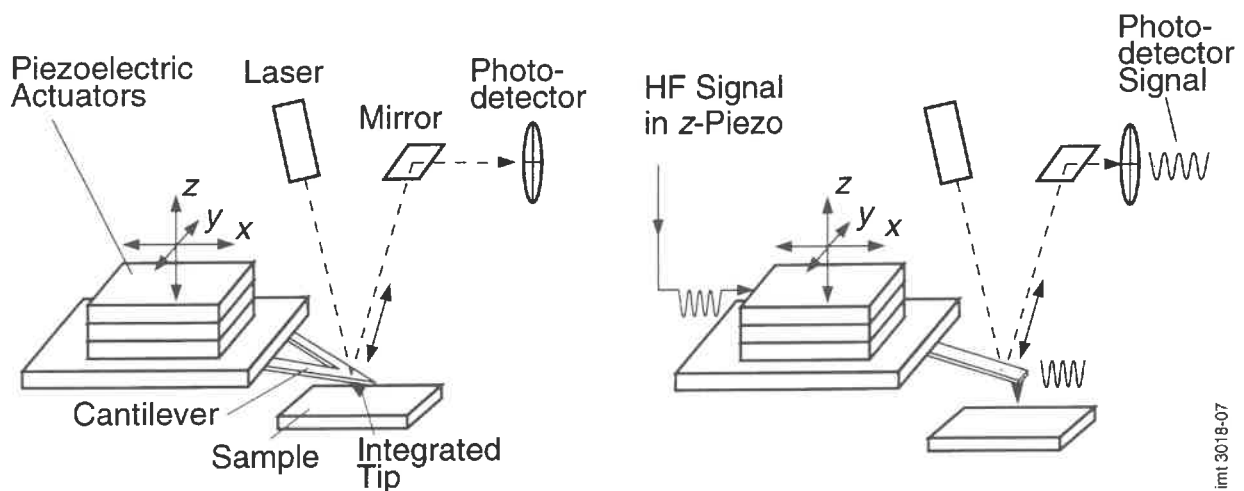


Fig. 2: Schematic of Atomic Force Microscopy (AFM)
 a. Contact (left)
 b. Non-contact (right)

is proportional to the probe's elevation at any location with the coordinates x and y , allowing a computer to establish a three-dimensional (3-D) surface topography plot. The resolution, which depends on the area measured, is in the range of smaller than 0.1 nm in z -direction and smaller than 1 nm in x - and y -direction.

While the contact mode measurement benefits from repelling forces between two bodies touching each other, the non-contact mode mostly exploits attraction between tip and probe surface due to surface absorption layers. There are also differences in picking up the surface topography's z -direction. The measurement takes advantage of the fact that a vibrating cantilever beam's resonance frequency changes when a force attacks at the non-supported end. For non-contact measurements (Fig. 2b.), a z -direction excitation causes the cantilever beam to vibrate at an oscillation frequency near its resonance frequency, which is approximately 100 kHz. When suffering a change in phase or amplitude, the z -direction is adjusted to restore the initial vibration condition [4]. The resolution in non-contact mode is comparable to the one in contact mode.

The stylus tip measurements were conducted with a TopoMetrix AFM, type TMX 2000, using a $100\ \mu\text{m}$ (x) $\times$ $100\ \mu\text{m}$ (y) $\times$ $10\ \mu\text{m}$ (z) scanning range. As mentioned above, the tips used in contact mode are $4\ \mu\text{m}$ long, the tip radius is smaller than 50 nm. For non-contact mode, the tips used are $8\ \mu\text{m}$ long and feature a tip radius smaller than 20 nm. In this case, the AFM tip radius is approximately a 100 times smaller than the profiler tip radius and thus very well suited to scan the profiler tip radius. Table I lists additional AFM cantilever and tip characteristics.

For AFM diamond stylus measurement, one of the main challenges is to design a proper fixture for the diamond stylus in order to find the profiler tip easily without damaging the AFM tip. Further, an accurate vertical position of the profiler tip in the fixture within $\pm 3\ \mu\text{m}$ has to be assured.

TABLE I
Contact and Non-Contact AFM Tips

	Contact	Non-Contact
Material	Silicon Nitride	Silicon
Geometry	Pyramidal	Triangular Pyramid
Aspect Ratio	1:1	3:1
Tip Radius	< 50 nm	< 20 nm
Cantilever Geometry	V	I
Cantilever Force Constant	0.032 N/m	24-85 N/m

VI. TEST RESULTS

For determining stylus geometry and in particular stylus radius, a SEM analysis (representing the state of the art in stylus analysis) and an AFM analysis were performed. Fig. 3 presents the SEM micrographs of a stylus with 90° cone angle and 5 μm radius. The top view of the stylus tip in Fig. 3a clearly shows artifacts apparently caused by the polishing process and fractures in the sphere region. The side view of the stylus tip in Fig. 3b allows to roughly estimate the stylus radius. The plane, in which the radius is shown, is not clearly defined since it is the stylus contour for the view given. The angle is distorted by the stylus' tilt inside the SEM and therefore the cone angle may not be determined. However, the micrograph allows a qualitative analysis of the stylus general shape and surface finish.

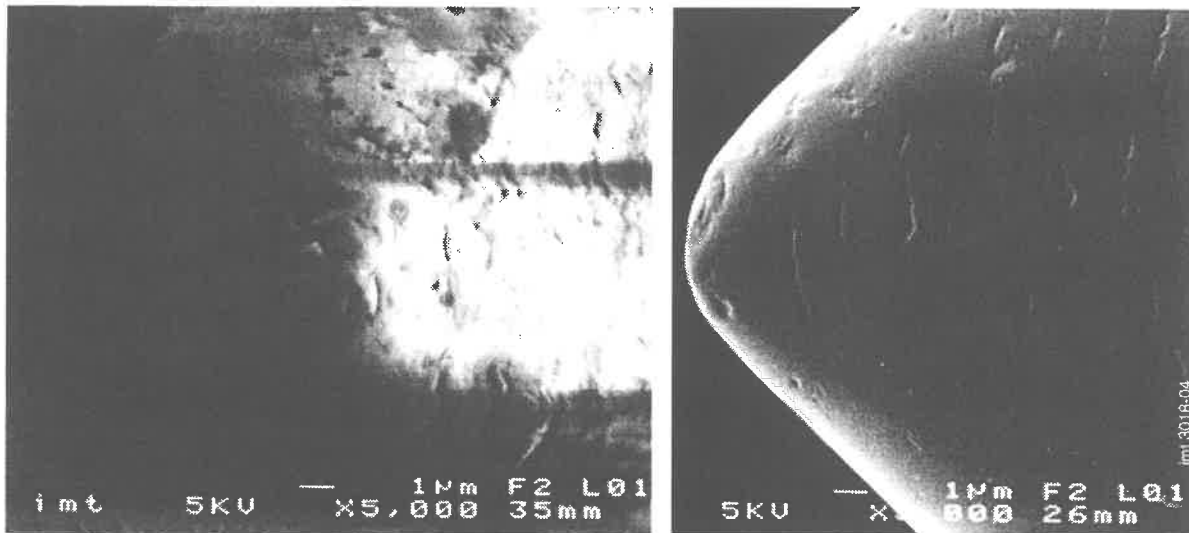


Fig. 3: SEM micrograph of a profiler stylus (test stylus A)
a. Top view (left)
b. Side view (right)

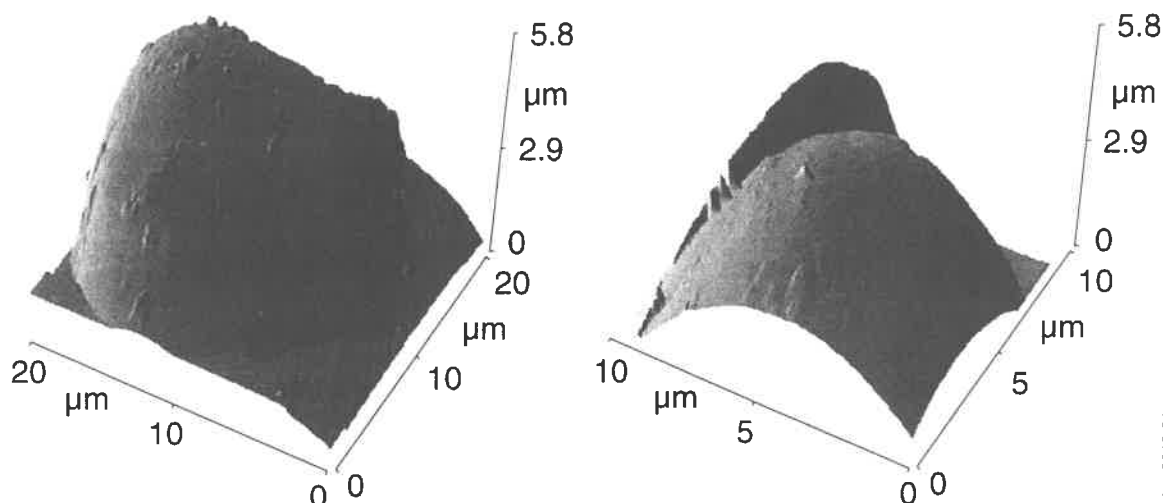


Fig. 4: Diamond stylus AFM analysis - topography
 a. Contact mode AFM measurement of test stylus A.
 b. Non-contact mode AFM measurement of test stylus B (reject - cracked)

Fig. 4a. depicts an AFM 3-D plot measured in the contact mode of the same stylus as shown in Fig. 3. As in case of the SEM analysis, the AFM plot shows polishing artifacts, although not quite as detailed as the SEM micrographs. However, the AFM provides for 3-D topography data up to a depth limited by the tip length. All 3-D topography data are stored, allowing to reconstruct any view of the specimen at any time.

Fig. 5 provides an AFM image showing the top view of a diamond tip with vertical and horizontal line scans. With the 2-D profiles running through the highest point of the tip, the coordinates of three points on the curve are sufficient to calculate the tip radius. The calculated radii are $6.5 \mu\text{m}$ for the horizontal profile and $4.9 \mu\text{m}$ for the vertical profile. The results correlate with the expected radius of $r = 5 \mu\text{m}$.

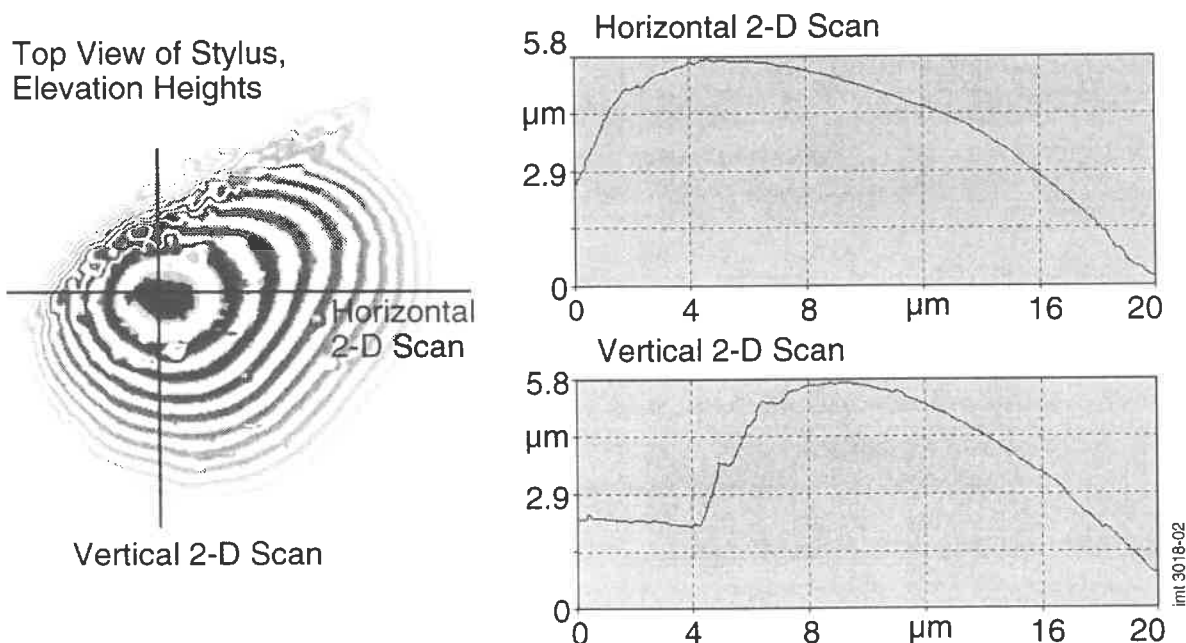


Fig. 5: Diamond stylus AFM analysis - top view and two line scans

Fig. 4b. depicts the AFM analysis in non-contact mode. The specimen is a fractured diamond stylus, the crack is clearly visible. Thus, both measurement modes, the contact as well as the non-contact mode, are very well suited to conduct diamond tip analysis. Furthermore, AFM diamond stylus analysis not only lends itself to establishing key stylus dimensions, it also allows to sort out defects.

V. CONCLUSION

AFM lends itself to profiler diamond stylus quality control, resolving the classic issue how to accurately measure the diamond stylus geometry and, most importantly, their radius. Both contact and non-contact modes are very well suited for the analysis. However, non-contact mode is superior, allowing to use longer and more pointed tips. Qualitative as well as quantitative analysis may be performed: Tip defects like fractures from the polishing process as well as the tip radius are obtained by the 3-D data set. Compared to the present SEM analysis, the AFM measurement is more time consuming, but contrary to SEM it allows to determine the stylus' radius.

ACKNOWLEDGEMENT

The authors are indebted to Mahr GmbH, Göttingen, Germany for providing the diamond stylus samples.

REFERENCES

- [1] ISO 1880: Instruments for the measurement of surface roughness by the profile method - Contact (stylus) instruments of progressive profile transformation - Profile recording instruments. International Organization for Standardization, Switzerland, 1979
- [2] ISO 4288: Geometrical Product Specifications (GPS) - Surface texture: Profile method - Rules and procedures for the assessment of surface texture. International Organization for Standardization, Switzerland, 1996
- [3] Bodschwinna, H., Hillmann, W.: Oberflächenmeßtechnik mit Tastschnittgeräten in der industriellen Praxis. Beuth Verlag, Berlin - Köln, 1992
- [4] TopoMetrix Technical Report: AFM Imaging Modes, TopoMetrix GmbH, Darmstadt 1996

INSTRUMENT FOR ON-LINE MONITORING OF SURFACE ROUGHNESS OF MACHINED SURFACES

J. Glenn Valliant
Optical Dimensions
Bozeman, Montana

Jean M. Bennett
Consultant
Ridgecrest, California

There is a need for on-line monitoring of surface roughness in the production environment. For example, in a mill where sheets of steel or aluminum are produced, the surface roughness on the rolls must be controlled to keep the roughness of the sheets within tolerance and control friction and speed of sheet metal in the rolling process. Monitoring the roughness along the length of a roll during the roll-grinding process makes it possible to achieve a much more uniform surface finish than is possible by the traditional monitoring of secondary parameters. On the production line, continuous monitoring of the surface roughness can detect changes of roughness and isolated defects so the roll can be changed before catastrophic failure occurs. Time and money are also saved by not having to regrind the roll surface until it is necessary.

Currently, surface roughness is measured during the roll-grinding operation by a portable mechanical profiler which is used to sample the roughness at a few places on the roll. One-hundred percent inspection of the roll surface is not possible with this type of instrument, and the measurements are off-line.

A new in-process roughness measurement instrument, the Lasercheck, has been developed to perform high speed, accurate, non-contact measurements of surface roughness. The instrument measures light scattered from the surface (Fig. 1). A visible diode laser emits a laser beam that passes through a window on the bottom of the gauge head and strikes the surface at a 75 degree angle of incidence. The light is reflected and scattered by the surface back into the gauge head where it falls on a photodiode detector array whose elements are arranged to measure the intensity of the specular beam and near-angle scattered light corresponding to surface spatial wavelengths ranging from greater than 0.5 mm to about 13 μm (see Table 1). Two isolated photodiodes detect scattered light at angles of about 37 degrees and 70 degrees from the specular direction, corresponding to surface spatial wavelengths of 1.9 μm and 0.8 μm , respectively. Thus, the photodiodes measure roughness contributions from surface spatial wavelengths in nearly the same bandwidth region that is measured by a mechanical profiler, the base-line instrument.

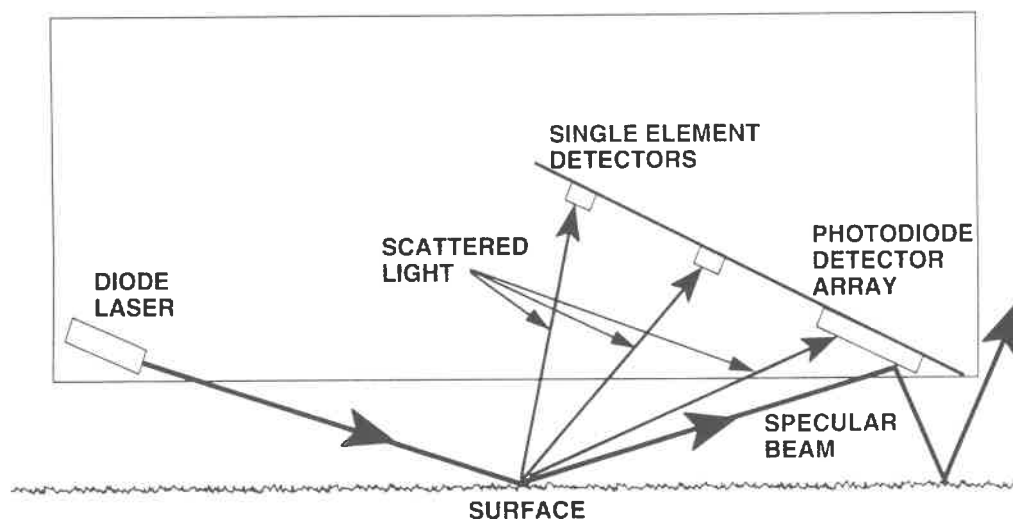


Fig. 1. Schematic diagram of Lasercheck instrument.

Table 1. Detector angles, surface spatial wavelengths and spatial frequencies for the Lasercheck roughness measurement instrument.

Detector Number	Degrees	Degrees From Specular	Surface Spatial Wavelengths (mm)	Surface Spatial Frequencies (1/mm)
0	75.00	0.00	Infinity	0.00
1	74.72	0.28	0.529	1.89
2	74.44	0.56	0.260	3.84
4	73.86	1.14	0.126	7.94
6	73.27	1.73	0.081	12.30
8	72.66	2.34	0.059	16.94
10	72.04	2.96	0.046	21.88
12	71.39	3.61	0.037	27.15
14	70.73	4.27	0.031	32.75
16	70.05	4.95	0.026	38.71
18	69.35	5.65	0.022	45.06
20	68.62	6.38	0.019	51.81
22	67.88	7.12	0.017	58.99
24	67.12	7.88	0.015	66.62
26	66.33	8.67	0.013	74.74
35	38.06	36.94	0.0019	520.93
36	5.45	69.55	0.0008	1299.34

The intensity and angular distribution of the specularly reflected and scattered light are measured, the data are digitized, and an average roughness (R_a) value is calculated for the surface at the position of the illuminated spot. This information is then sent to a host computer where the R_a values are displayed on a graph, or are available as a data file. Data are provided 10 times per second, and each data point is an average of about 6 surface measurements.

Alignment checks are performed during each measurement cycle. The gauge head has been designed for a height of 1.06 ± 0.06 inches above the surface to be measured. Motions and vibrations within that tolerance range are monitored continuously, and reflectance and scatter distributions are normalized and corrected during every measurement cycle to ensure accurate results. If the instrument senses that the height has moved outside of the tolerance range, an error message appears on the monitor screen and the user is prompted to correct the problem. This feature prevents incorrect data from being incorporated into the measurement sequence. At the end of a group of measurements, the user can save the data for subsequent viewing or in one of a variety of file formats. The data can also be plotted directly from the screen or the R_a values can be printed.

The instrument is optimized to measure surface roughness in the $0.2 \mu\text{in}$ to $20 \mu\text{in}$ ($0.005 \mu\text{m}$ to $0.5 \mu\text{m}$) average roughness range. Reliable results can be obtained to roughness values $> 40 \mu\text{in}$ ($1 \mu\text{m}$) with careful alignment. A nominal $9.5 \mu\text{in}$ R_a ground steel surface will be used as an example of the type of data that can be obtained with the instrument. (Data from other samples will also be shown in the talk.) Figure 2 shows a Nomarski micrograph of the $9.5 \mu\text{in}$ surface. A profile of this surface taken with a Talystep surface profiler¹ is shown in Fig. 3. The measured roughness was $12.21 \pm 0.34 \mu\text{in}$ rms for an average of 10 places on the surface. The software for this profiler does not calculate R_a values; they will be slightly smaller than the rms roughness for this surface. Figure 4 shows a graph of intensity versus scattering angle (measured from the specular direction) taken with the Lasercheck instrument. The calculated R_a value from this data is $12.72 \mu\text{in}$. Other

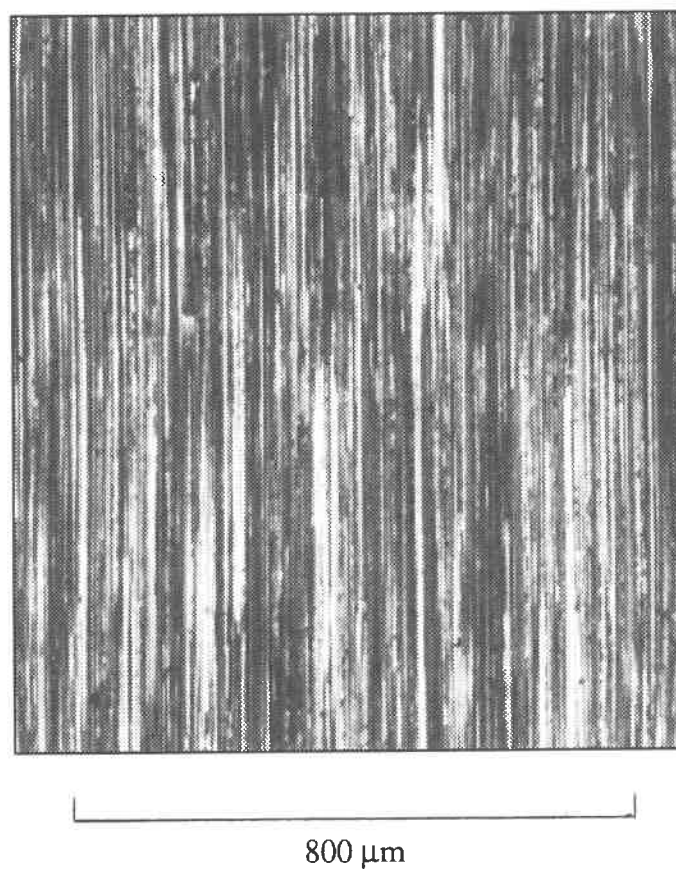


Fig. 2. Nomarski micrograph of nominal 9.5 μin *Ra* ground steel surface

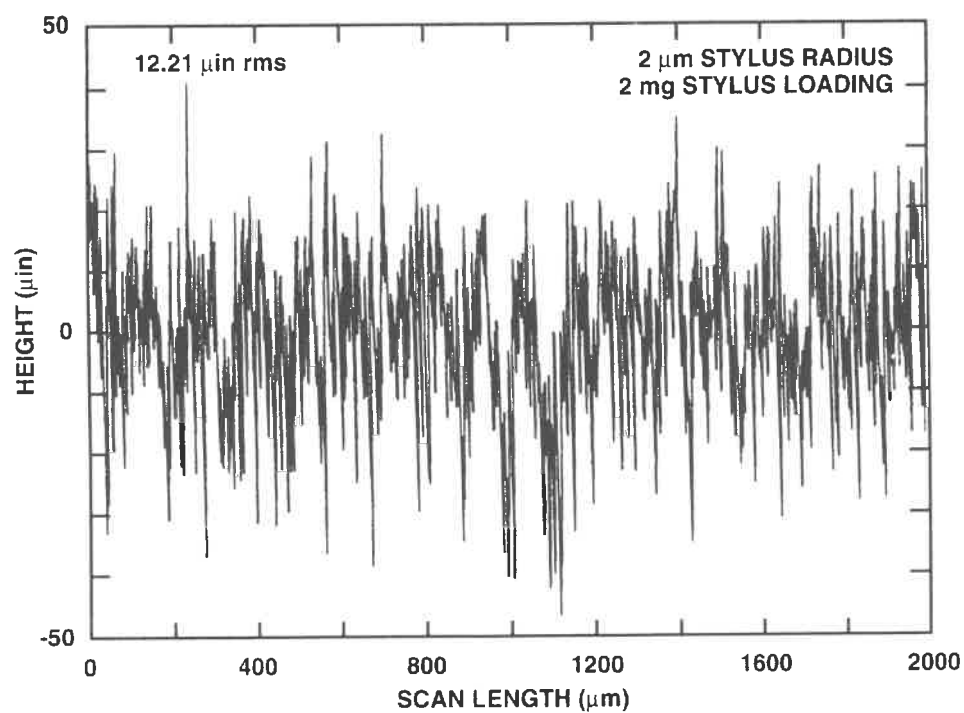


Fig. 3. Talystep surface profile of nominal 9.5 μin *Ra* ground steel surface.

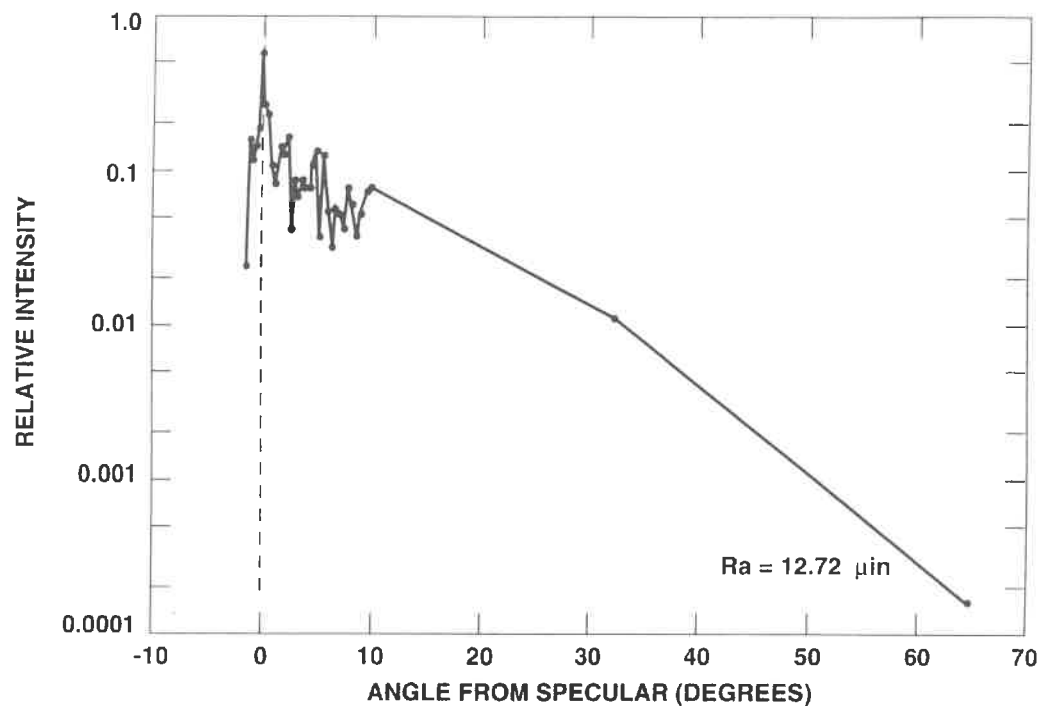


Fig. 4. Intensity versus scattering angle measured with the Lasercheck roughness measurement instrument.

Table 2. Roughness of ground steel surfaces (μin).

Nominal Roughness (Ra)	Lasercheck (Ra)	Taylor Hobson Surtronic 3P (Ra)	Taylor Hobson Talystep (rms)
1.5	1.83	2.56	1.66
9.5	12.72	10.62	12.21
35	34.56	30.92	35.55

ground steel surfaces in the roughness range from 1.5 μin to 35 μin showed good agreement between Talystep profile roughness values and those measured by the Lasercheck instrument, as shown in Table 2.

Figure 5 shows an example of the use of the instrument in a roll grinding operation. An ideal roll roughness map (over an approximately 6 foot roll length) is shown in Fig. 5(a). This surface was actually produced on a commercial roll. Figure 5(b) shows a more typical surface that is smoother in the middle and rougher at the ends, while Fig. 5(c) shows the type of surface that can be obtained during too aggressive grinding.²

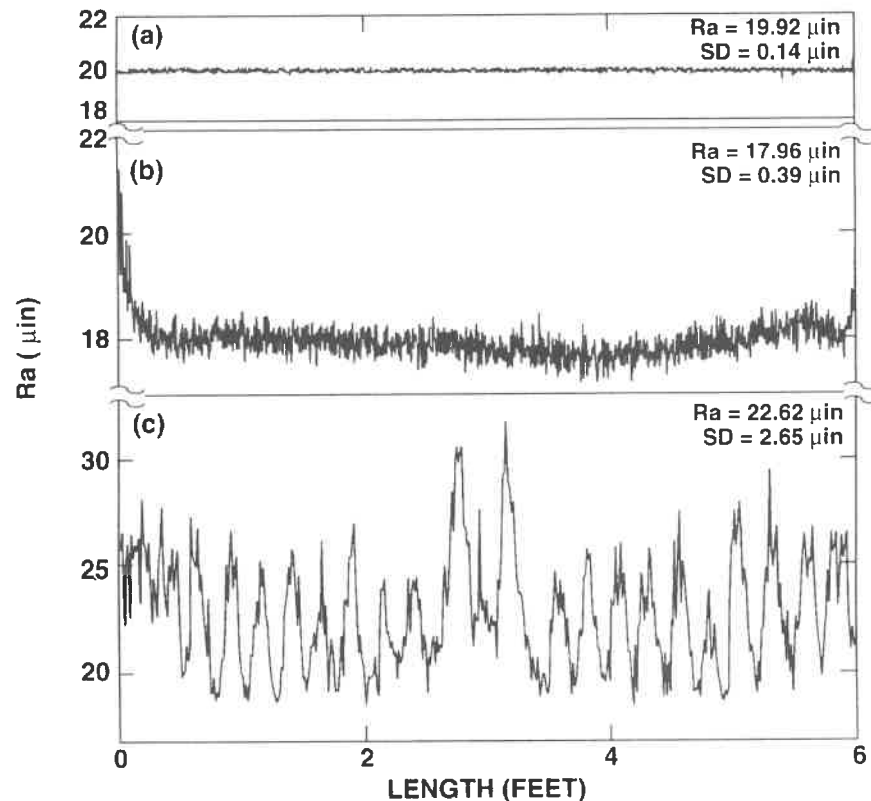


Fig. 5. Measured roughness profiles of rolls in a production environment: (a) ideal roughness profile, (b) typical roughness profile that is smoother in the center and rougher on the ends, and (c) oscillating roughness profile caused by too aggressive grinding.

Although the examples shown above are all for one-dimensional ground surfaces, the Lasercheck instrument can also be used to monitor surfaces made by other processes, for example by electrical discharge texturing, belt sanding, lapping and polishing, or surfaces on rollers used in the paper industry. The roughness magnitudes on these other types of surfaces are similar, so the instrument can easily be adapted for their measurement. When surface roughness is directly monitored and used to control the roll surfacing operation in a closed loop feedback circuit, the resulting surface roughness can be made much more uniform than when other parameters such as current from the grinding wheel motor, roll revolutions per minute, and feed rates are used. This closed loop technology is currently being implemented.

In summary, an instrument designed to be integrated into a production facility to measure surface roughness in the 0.2 μin to 40 μin *Ra* range has been described. The Lasercheck makes measurements at the rate of 0.1 second per measurement on-line and automatically corrects for sample vibration. Because of the possibility of close to 100% surface inspection and a repeatability of $\pm 0.1\%$, process-induced variations can be revealed and corrected. The instrument can be used for processes involving metal coils and machined metal surfaces, magnetic media, plastics, or ceramics. Calibration is traceable to the National Institute of Standards and Technology with an accuracy of $\pm 2\%$. The instrument is easy to use, non-contact, and hence does not damage the

surface. It will be finding uses in a variety of industries and will make possible considerable improvement in product surface finish uniformity, surface finish tolerances, less wasted material, and more cost effective use of production machinery.

References

- 1 Jean M. Bennett and Lars Mattsson, Introduction to Surface Roughness and Scattering (Optical Society of America, Washington, D. C., 1989) and references therein.
- 2 J. Glenn Valliant, "Detection of Roll Surface Roughness Defects Caused by the Grinder," (38th Mechanical Working and Steel Processing Conference, Cleveland, OH, October 13-16, 1996, Iron & Steel Society).

Author Index

<i>Badami, Vivek</i>	*	<i>Li, Zhu</i>	26, 42
<i>Bennett, Jean M.</i>	82	<i>Liu, Xiangyang</i>	10
<i>Blunt, Liam</i>	20	<i>Liu, Xiao Ming</i>	46
<i>Bowen, D. Keith</i>	*	<i>Maetzig, J. Chris</i>	76
<i>Brown, Christopher A.</i>	30	<i>Malburg, Mark</i>	36
<i>Buehring, I. K.</i>	51	<i>Mansfield, Daniel I.</i>	51
<i>Caber, Paul J.</i>	61	<i>Mooney, Charles B.</i>	*
<i>Cai, Guang Qi</i>	46	<i>Raja, Jay</i>	10
<i>Chen, Qinghu</i>	26	<i>Renegar, T. B.</i>	69
<i>Cohen, Donald K.</i>	30	<i>Roth, John</i>	57
<i>de Groot, Peter</i>	57	<i>Russell, Phillip</i>	*
<i>Dixson, Ronald G.</i>	69, 70	<i>Schmit, Joanna</i>	61
<i>Doi, T.</i>	69	<i>Siekmann, Hieko</i>	76
<i>Feng, Xiao Lin</i>	46	<i>Smith, Stuart T.</i>	3
<i>Fu, J.</i>	69, 70	<i>Song, Jun Feng</i>	69
<i>Gatzen, Hans H.</i>	76	<i>Stout, Kenneth J.</i>	20
<i>Hara, Seiichiro</i>	15	<i>Tsai, Vincent W.</i>	69, 70
<i>Harb, Salam M.</i>	3	<i>Valliant, J. Glenn</i>	82
<i>Hoch, Pierre</i>	30	<i>Vorburger, Ted V.</i>	69, 70
<i>Howard, Lowell P.</i>	*	<i>Wang, X. S.</i>	70
<i>Jarausach, Konrad F.</i>	*	<i>Williams, E. D.</i>	70
<i>Jennings, Richard</i>	65	<i>Wyant, James C.</i>	61
<i>Jiang, Xangquian (Jane)</i>	20, 26, 42	<i>Xiao, Shaojun</i>	42
<i>Johnsen, William A.</i>	30	<i>Xie, Tiebang</i>	42
<i>Kobayashi, Yoshikazu</i>	15	<i>Yanagi, Kazuhisa</i>	15
<i>Köning, R.</i>	70	<i>Yang, Shunian</i>	26

* Abstract not available at press time

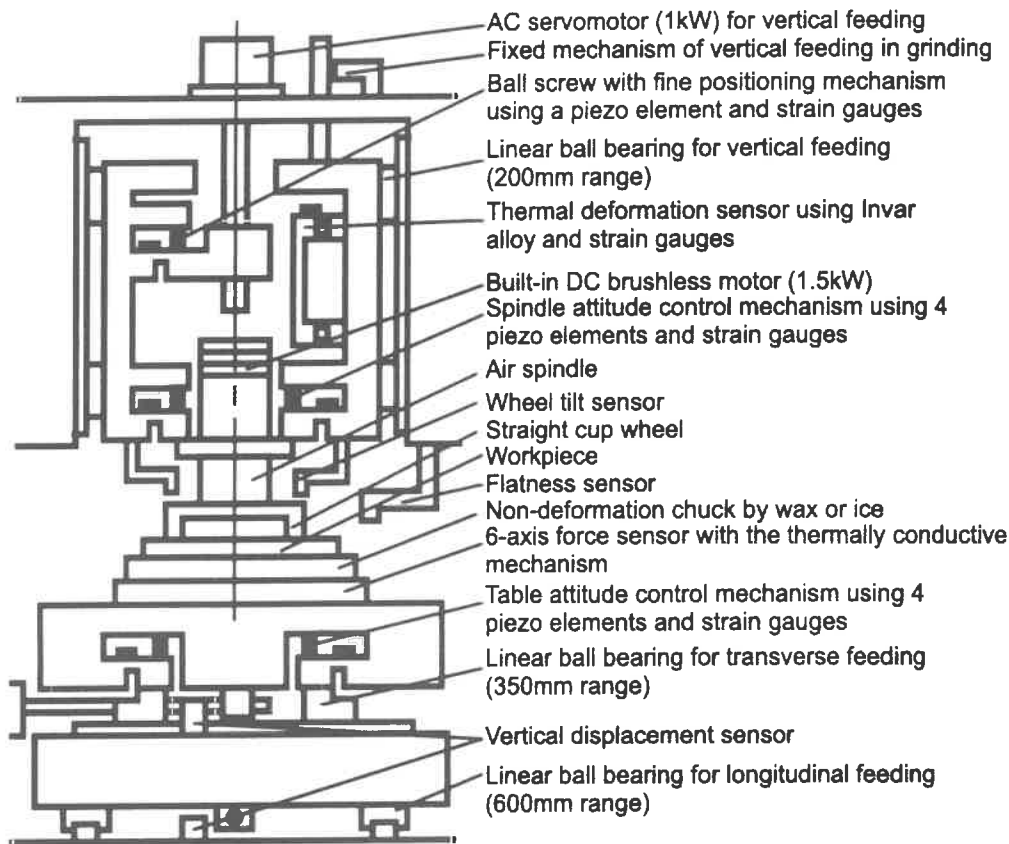


Fig. 2 a conceptual sketch of the self-deformable face grinding machine

terial ; the metal machine expands as the temperature increases, but the sensor doesn't ; strain gauges on the sensor measure the deformation caused by the discrepancy of the thermal expansions. The sensor can detect the expansion with a $0.1 \mu\text{m}$ resolution per 1,300 mm of length of the legs. The wheel tilt sensor is composed of four capacitance displacement sensors. The gap between the wheel and the sensor is always free from debris or mists because the air flow exhausted from the spindle works as an air seal. The sensor can detect Z height with a 10 nm resolution and tilt with $0.03 \mu\text{rad}$.

4. Evaluation of self-deformable face grinding machine

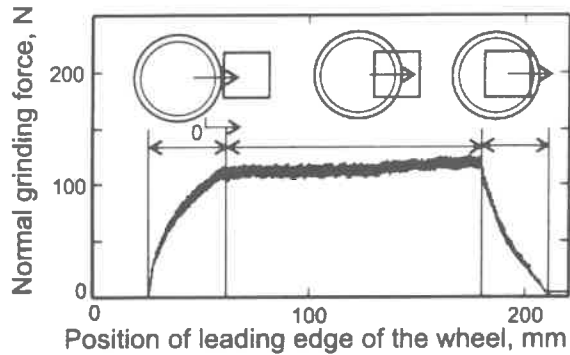
4.1 Force-induced deformation compensation In order to produce ultra-flat surfaces using a straight cup wheel in face grinding, it is most important to keep the wheel surface normal to the spindle axis. Otherwise the tilted wheel makes a curved surface on the workpiece.

Figure 3 shows normal grinding force history ; the force history indicates that the upward pushing force in the wheel initially increases as the wheel engages the workpiece, keeps constant, and decreases toward the end ; Figure 4 shows that the force pushed the wheel upward to $3 \mu\text{m}$ and the moment declined it front-upward to 12

μrad . Figure 5 shows ground surface profiles : convex curvatures of $2.5 \mu\text{m}$ in the transverse feeding (Y) direction and those of $2.0 \mu\text{m}$ in the longitudinal feeding (X) direction. The profiles were measured while they were still fixed on the table. The high rigidity of the machine, which is high in Z direction to $70 \text{ N}/\mu\text{m}$ but is not so high in Y rotation to $2 \text{ Nm}/\mu\text{rad}$, improves surface flatness from about $10 \mu\text{m}$ to $2.5 \mu\text{m}$. But only the passive mechanisms can't make some $0.1 \mu\text{m}$ flatness.

Next we kept the wheel flat by the wheel tilt sensor and the spindle attitude control mechanism. Figure 6 shows the attitude of the wheel under the control ; the wheel kept to the initial Z position within $\pm 0.5 \mu\text{m}$ and kept flat within $\pm 2 \mu\text{rad}$. Figure 7 shows surface profiles ground under the control : convex curvatures of $0.8 \mu\text{m}$ in the Y direction and $0.6 \mu\text{m}$ in the X. Through our grinding experiences, 8-inch size substrates can be produced with a $0.5 \mu\text{m}$ flatness except a 5mm-wide edge area.

4.2 Thermal deformation compensation The wheel tilts when the column deforms irregularly. For example, under room temperature variation of 16 degrees, the front side legs expand by about $70 \mu\text{m}$ more than the back



wheel : SD #325 bronze bonded (ϕ 220x w5 mm), spindle rotation : 1500r.p.m, depth of cut : $5\text{ }\mu\text{m}$, workpiece : glass (150x150x7mm), feed rate : 200mm/min, grinding coolant : water soluble type flowed from the spindle center.

Fig. 3 Normal grinding force in grinding without the wheel attitude control

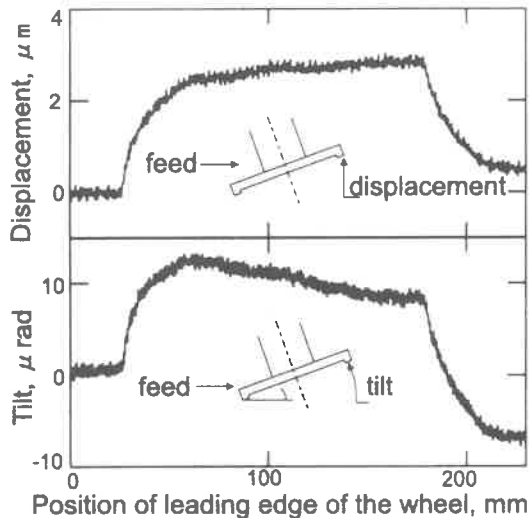


Fig. 4 Displacement and tilt of the wheel in grinding without any control

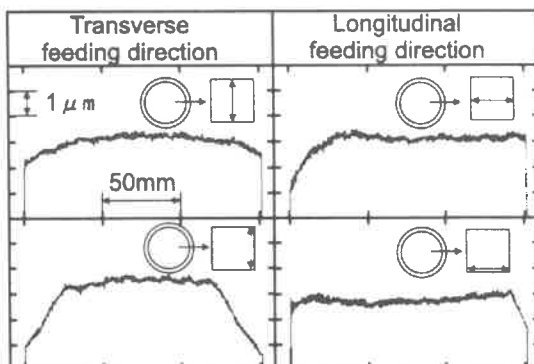


Fig. 5 Profile of the ground workpiece in grinding without any control

side legs : the wheel translates upward by $14\text{ }\mu\text{m}$ and tilts front-upward by $16\text{ }\mu\text{rad}$, which is comparable to tilts in grinding shown in figure 4. In addition to room temperature variation, the motor coolant flow variation induced $2\text{ }\mu\text{rad}$ and the spindle rotation did $4\text{ }\mu\text{rad}$. We would suffer various thermal deformations ; only a $5\text{ }\mu\text{rad}$ tilt of the ϕ 220 mm wheel would degrade flatness by $0.5\text{ }\mu\text{m}$.

To compensate for the deformation, heaters on the column legs (see Figure 1) was effective in addition to the spindle attitude control mechanism [1]. Figure 8 shows that the wheel tilts by $5\text{ }\mu\text{rad}$ when the heaters on two front (or back) side legs expand by $30\text{ }\mu\text{m}$.

But a passive method is more cost-effective than an active method. Figure 9 shows the difference of the thermal expansions of the legs without / with the thermally insensitive surface made of a 10 mm thick foamed plastic. Figure 9 indicates that the insensitive surface improved the thermal expansion to only about 30 %. We

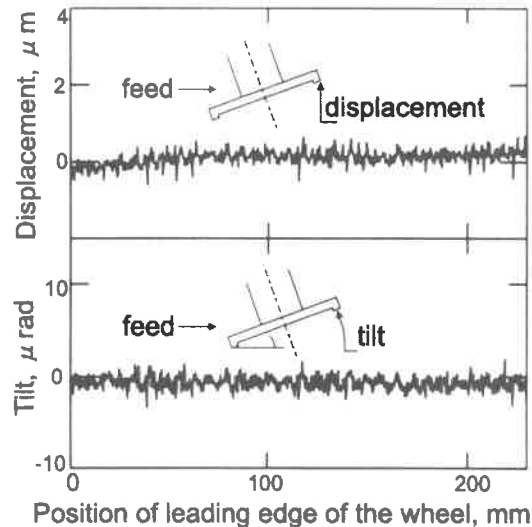


Fig. 6 Displacement and tilt of the wheel in grinding with the wheel attitude control

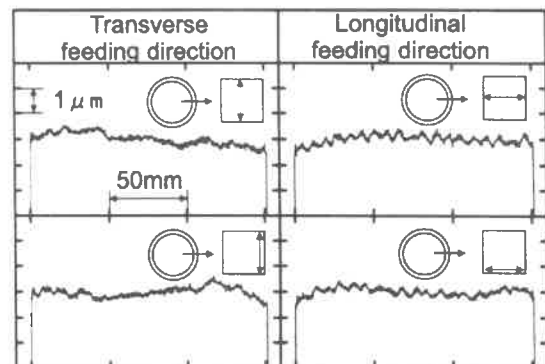


Fig. 7 Profile of the ground workpiece in grinding with the wheel attitude control

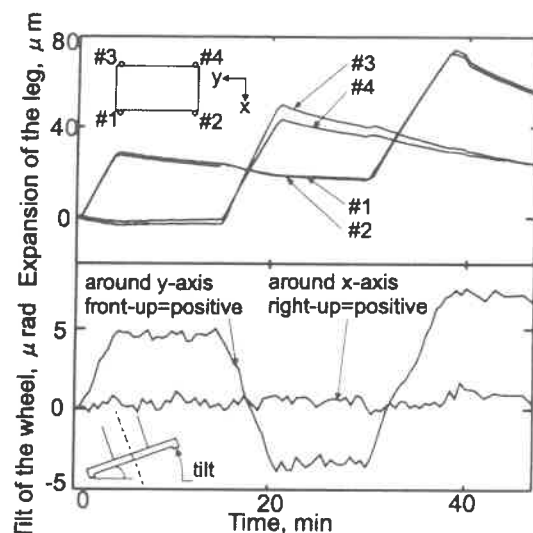


Fig. 8 Tilt of the wheel induced by expansion of the legs of the column

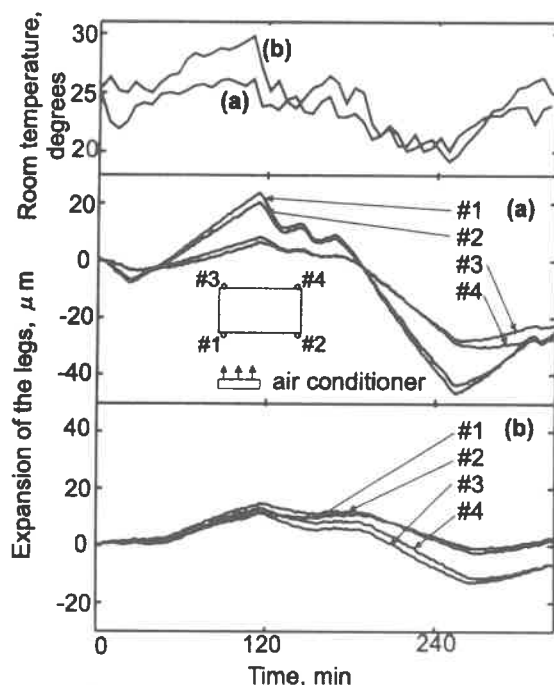


Fig. 9 The thermal expansion of the legs (a) without / (b) with thermally insensitive surface

can neglect the thermal deformation of the column with the thermally insensitive surface.

4.3 Bearing fluctuation compensation The flatness is also affected by a Z height deviation caused by the linear ball bearing of the table. Figure 10 shows the non-compensated Z height deviations on the top and the compensated ones on the bottom. Each shows five measured deviations. The former indicates that the waviness had

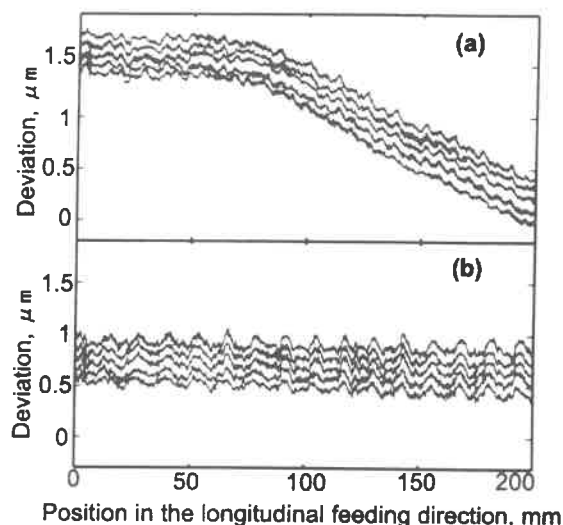


Fig. 10 (a) Z height deviations measured five times that are caused by the linear ball bearing of the table, (b) the compensated ones with the table attitude mechanism

large repeatability, but the roughness looked like spikes did small repeatability. Because the interval of the spikes is twice of the ball diameter, the roughness should be caused by irregularity of ball movement from the loose return groove to the tight preloaded slider groove. When some spikes in the former and no spikes in the latter happen unfortunately, the compensated roughness made worse as indicated in Figure 10.

The measured Z height deviation over a 200 x 200 mm area is about 1.5 μm . We can compensate it using the table attitude mechanism within $\pm 0.05 \mu\text{m}$ the waviness and within $\pm 0.1 \mu\text{m}$ the roughness.

5. Conclusion

To produce some 0.1 μm flatness on the 8-inch substrates of information processing devices, we designed, fabricated, and evaluated a "self-deformable face grinding machine". The machine executes an active, knowledge-based, self-deformable control using sensors, actuators and computers. The machine improves the flatness of ground substrate surfaces up to 0.5 μm under the control. Further improvement from 0.5 μm to 0.1 μm will be realized by some developing mechanisms which weren't described in this paper : a ball bearing with a retainer or a zero-deformation chuck with ice.

We thank THK Co., Ltd. for fabricating the machine.

Reference

- [1] Hatamura, Y., et al., CIRP Ann. 42-1, 549-552, 1993
- [2] Nakao, M., et al., CIRP Ann. '96, 397-400, 1996
- [3] Matsumoto, K., et al., ASPE Ann. '96, 476-481, 1996

A Local Grinding System for Wide and Flat Surface with a 3-Axis Negative Compliance Device

Kiyoshi MATSUMOTO, Masayuki NAKAO and Yotaro HATAMURA
Department of Engineering Synthesis, Faculty of Engineering, The University of Tokyo
7-3-1 Hongo Bunkyo-ku Tokyo 113, JAPAN

1. Introduction

Machining process to obtain a wide and flat surface in short time is highly required in many industrial fields; silicon wafers, magnetic disk substrates, ceramic substrates for computers and optical flats. Usually whole lapping method is employed for this processing, but it is inefficient from view points of process time and material cost. The authors challenge local grinding to obtain a wide and flat surface in short time.

Generally, to increase preciseness of tool position control, rigidities of frame, chuck and tool of a grinding system have to be reinforced. This causes increases of size, mass and complexity of the system. The authors make an another approach. By employing a negative compliance device, it is possible to compensate deformation of the grinding system caused by machining force.

The authors previously proposed a force sensor with actively controlled compliance.[1] With this sensor, it is possible to achieve negative compliance. We applied this sensor to a local lapping/grinding system with a water driven turbine motor.[2] Through the experiment of grinding a silicon wafer, it was confirmed that grinding error was reduced to 60 %. But there was a deviation in distribution of the grinding force on the tool. To increase preciseness of tool position control, it is necessary to control the tool not only in one direction but also in two moment-directions. The authors developed an improved 3-axes negative compliance system, and applied it to a local grinding system.

2. Analysis of grinding process

Fig. 1(a) shows schematic of conventional grinding process. If a grinding tool is controlled so as to follow a prescribed trajectory, the resulting surface does not retain the desired accuracy. The grinding force deforms the mechanical elements such as the tool, the workpiece and the structural frame. This results in grinding error by spring-back.

Fig. 1(b) shows schematic of grinding process with negative compliance. To compensate spring-back, a negative compliance device generates additional depth of cut. When initial depth of cut is Z , grinding force F , spring-back S and grinding error E are described as follows.

$$F_0 = -\frac{Z}{C_d} \quad (1), \quad S_0 = C_s \cdot F_0 \quad (2), \quad E_0 = S_0 \quad (3),$$

C_d is dynamic compliance, and defined as a ratio of (grinding force)/(cutting depth). C_s is static compliance

of the grinding system. When C_z is employed as compliance to generate additional depth of cut Z_c , grinding force F , spring-back S and grinding error E are given as follows.

$$Z_{c1} = -C_z \cdot F_0 \quad (4), \quad F_1 = \frac{Z + Z_{c1}}{C_d} = \frac{Z - C_z \cdot F_0}{C_d} \quad (5),$$

$$S_1 = C_s \cdot F_1 \quad (6), \quad E_1 = S_1 - Z_{c1} \quad (7),$$

F_i is a geometric series. If C_z is within the range of $-C_d < C_z < C_d$, grinding force F_i converges to a certain value, and F , S , Z_c and E are given as follows.

$$F_\infty = \frac{Z}{C_d + C_z} \quad (8), \quad S_\infty = C_s \cdot F_\infty \quad (9), \quad Z_{c\infty} = -C_z \cdot F_\infty \quad (10),$$

$$E_\infty = S_\infty - Z_{c\infty} = (C_s + C_z) \cdot F_\infty = \frac{C_s + C_z}{C_d + C_z} \cdot Z \quad (11),$$

Equation (11) shows that the grinding error can be zero regardless of the cutting depth Z if C_z is set to $-C_s$. This means that if the static compliance of the grinding system

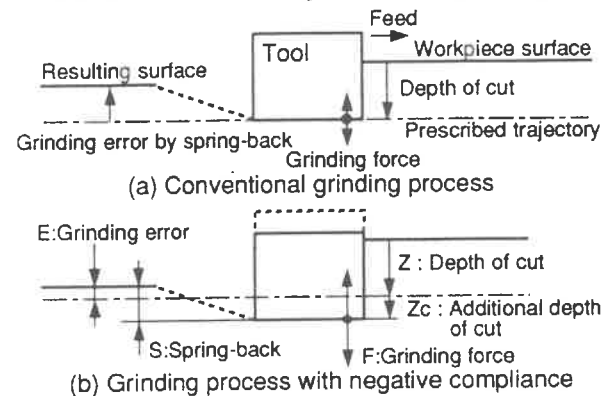


Fig.1 Schematic of grinding process

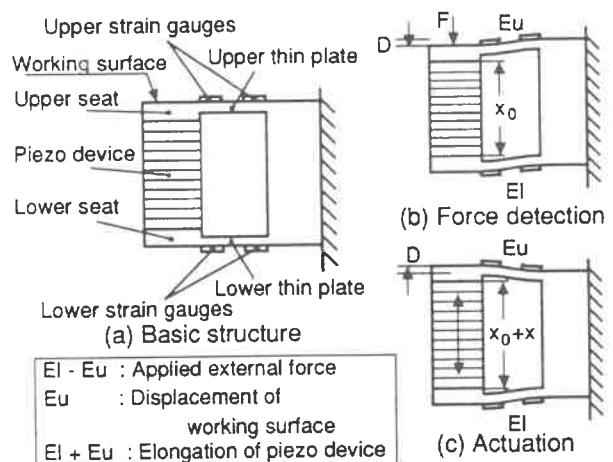


Fig.2 Structure and deformations of sensing unit

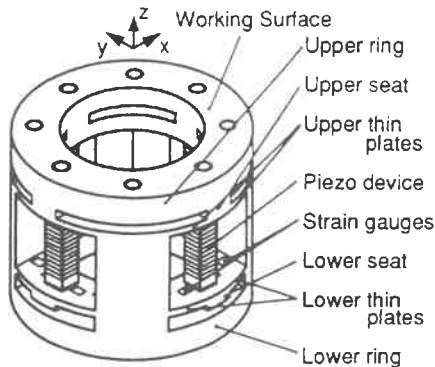


Fig.3 Ring-shape active force sensor

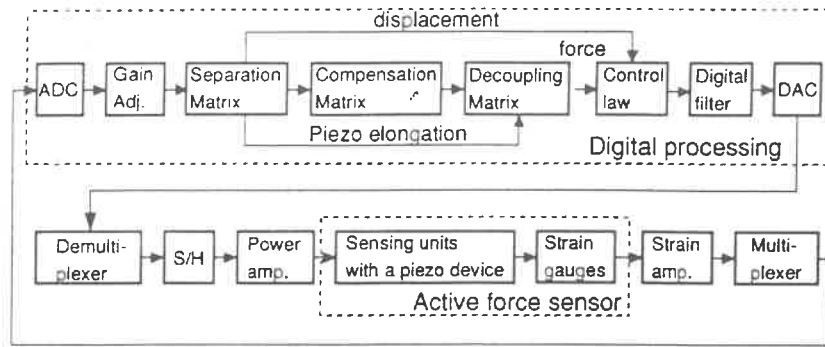


Fig.4 Block diagram of sensing and actuating system

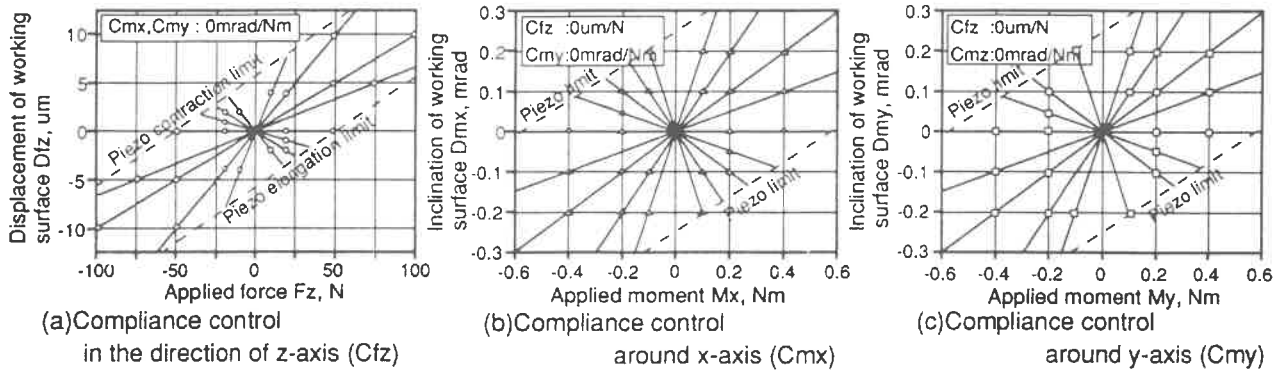


Fig.5 Compliance control operation

Cs is measured precisely, an error free grinding system can be achieved.

3. Development of a 3-Axes negative compliance system

3.1 Structure of a ring-shape active force sensor

This Fig. 2(a) shows structure and deformations of a sensing unit. A pillar is composed of upper/lower seats, and a piezo device which is glued between the seats. This pillar is movable and connected to a fixed pillar by a pair of parallel thin plates. On the surface of the thin plates, strain gauges are glued which detect their deformations.

External force applied onto the working surface causes displacement (D_f) as shown in Fig2(b), and it is detected by the strain gauges. Elongation of the piezo device generates displacement of working surface (D_a) as in Fig2(c), and it is detected by the strain gauges similarly. The active force sensor uses superposition of these deformations. Assuming that the signs of E_u and E_l are positive as shown in the case of Fig. 2(c). E_u is proportional to the applied external force, and is not affected by the elongation of the piezo device. E_u , $E_u + E_l$ are proportional to the displacement of the working surface and the elongation of the piezo device respectively.

Fig. 3 shows structure of the ring-shape active force sensor. Between upper and lower ring, 4 active sensing units are located at every 90 degrees. Each unit can detect force and actuate. By calculation of the outputs from 4 units,

force in the direction of z (F_z) and moments around the x and y axes (M_x , M_y) can be obtained. By actuation of the piezo elements in 4 units, displacement in the direction of z (D_{fz}) and inclinations around the x and y axes (D_{mx} , D_{my}) can be generated. The lower ring is fixed to frame or base, and a machining tool is attached to the upper ring.

3.2 Sensing and actuating system

Fig. 4 shows a block diagram of the sensing and actuating system. The output signals from strain amplifiers are converted to digital form, being processed in a computer. Control signals are converted to analog form, and drives the piezo devices in the sensing units through power amplifiers. In digital processing session, the separation matrix is used to separate force, displacement and piezo elongation information. The compensation matrix is used to reduce mutual interference with force information F_z , M_x and M_y . The decoupling matrix removes interference with force information due to actuation.

Fig. 5 shows performance of compliance control operations. The compliance can be changed within the working range of the piezo devices. As shown in Fig5(a), the rated value of force F_z is about $-50\text{N} \sim 50\text{N}$ when compliance C_{fz} is set to zero. As shown in Fig5(b) and (c), the rated value of moments M_x and M_y are about $-0.6\text{Nm} \sim 0.6\text{Nm}$ when compliance C_{mx} and C_{my} is set to zero. The resolutions of force and displacement information

were estimated to be less than 0.5%.

4. Grinding experiment

4.1 Experimental system

Fig. 6 shows the experimental system. A surface grinding machine is used as feed mechanism. Straight cup wheel is employed as grinding tool, and is rotated by a water driven turbine. The pressure of water fed to water inlet is about 3 MPa, and nominal rotation speed is 10000 r.p.m. Rotation shaft which connects the turbine and the grinding tool is suspended by a dynamic water pressure bearing. The ring-shape active force sensor is installed between the turbine and the feed mechanism in z direction. An 8-inch silicon wafer is glued to the table base by hot wax. The table base is fed in x direction. The static compliance C_s of this system is measured to be $0.257 \mu\text{m/N}$ at the center of the table base.

Fig. 7 shows nominal experimental conditions. Diameter of the straight cup wheel is 30mm. Metal bonded grains on the wheel is #2000. Feed rate is fixed at 100 mm/min. Bandwidth of the digital servo system to control the compliance is set at about 10 Hz. DC gain is about 60 dB. Grinding forces F_z , M_x , M_y and displacements of the sensor D_{fz} , D_{mx} , D_{my} are measured when grinding over a trapezoid shape. Nominal height of the trapezoid shape is $10 \mu\text{m}$, length is 80mm. As for attitude of the wheel, the inclination and the compliance around y-axis (D_{my} and C_{my}) are fixed at -0.1 mrad and 0 mrad/Nm respectively, so that leading edge of the wheel can always be used for cutting.

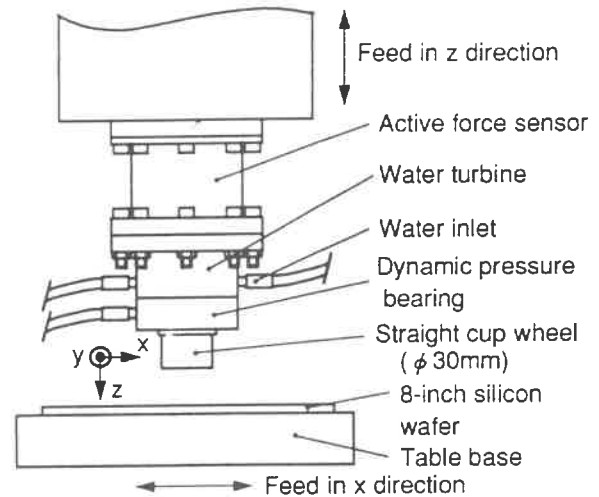


Fig.6 Experimental system

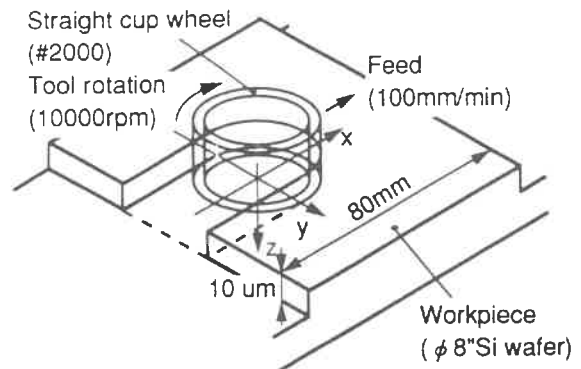


Fig.7 Experimental conditions

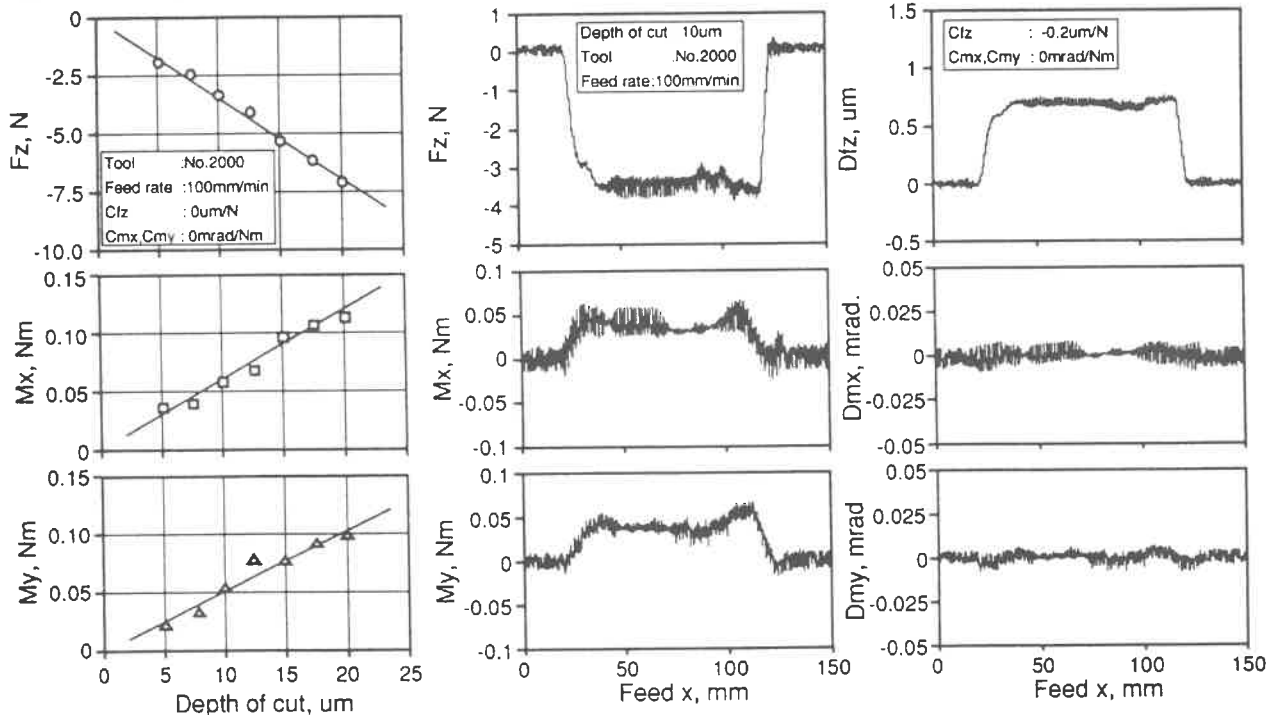


Fig.8 Depth of cut vs. grinding force Fig.9 Grinding forces and displacements grinding over a 10μm-step

4.2 Measurement of the dynamic compliance C_d

Grinding forces are measured under the condition that all the compliance is set to zero. Fig. 8 shows cutting depth vs. grinding forces. The forces are in proportion to the cutting depth. The dynamic compliance in z-axis C_d is estimated to be $2.80\mu\text{m/N}$. The dynamic compliance C_d is 10 times greater than the static compliance C_s .

4.3 Grinding experiment with negative compliance

In this experiment, we measured reactive force when grinding over a trapezoid shape of $10\mu\text{m}$ high. Fig. 9 shows an example of grinding forces and displacements. The compliance C_{fz} is set at $-0.2\mu\text{m/N}$. The grinding force F_z is about 3.5N , and the sensor generates deformation of $0.7\mu\text{m}$ to reduce the grinding error. The inclination around x-, y-axis D_{mx} , D_{my} are almost zero.

Due to the analysis in section 2, the grinding error will be zero when the compliance C_{fz} is set at $-C_s$, i.e. $-0.257\mu\text{m/N}$. Fig. 10 shows compliance C_{fz} vs. grinding error under the condition that compliance C_{mx} and C_{my} is set to zero. The error is minimum when the compliance C_{fz} is $-0.3\mu\text{m/N}$. There is deviation in the error according to the position y. The error at the left edge of feed path is $0.2\mu\text{m}$ higher, while that at the right edge is $0.2\mu\text{m}$ lower. This derives from deviation in distribution of the grinding force. The moment M_x represents this deviation, so negative compliance around x-axis is introduced.

Fig. 11 shows compliance C_{mx} vs. grinding error under the condition that the compliance C_{fz} is set to $-0.3\mu\text{m/N}$, and C_{my} to zero. The error is minimum when the compliance C_{mx} is -0.3mrad/Nm . The error is less than $0.1\mu\text{m}$.

Fig. 12 shows resulting surface profile when the compliance C_{fz} is changed. The compliance C_{mx} and C_{my} is zero. Length of a trapezoid shape is 30mm . Grinding error increases at the end of feed. It probably derives from change of the static compliance C_s at edge of the base.

5. Conclusions.

A local grinding system with a 3-axis negative compliance device was proposed. By analysis of grinding process, the condition to minimize grinding error was shown. 3-Axes negative compliance system with a ring-shape active force sensor was developed, and its performance was confirmed. Through experiment of grinding a silicon wafer, it was confirmed that grinding error was reduced to $0.1\mu\text{m}$.

[1] K. Matsumoto, et al., A Trial of Force Sensor with actively controlled Compliance, Proc. ASPE Spring Topical Meeting on Mechanisms and Controls for Ultraprecision Motion, ASPE, Tucson, USA, (1994)

[2] K. Matsumoto, et al., Local Lapping System to Obtain Wide and Flat Surface with Actively Controlled Force

Sensor, Proc. 1996 ASPE Annual Meeting, ASPE, Monterey, USA, (1996)

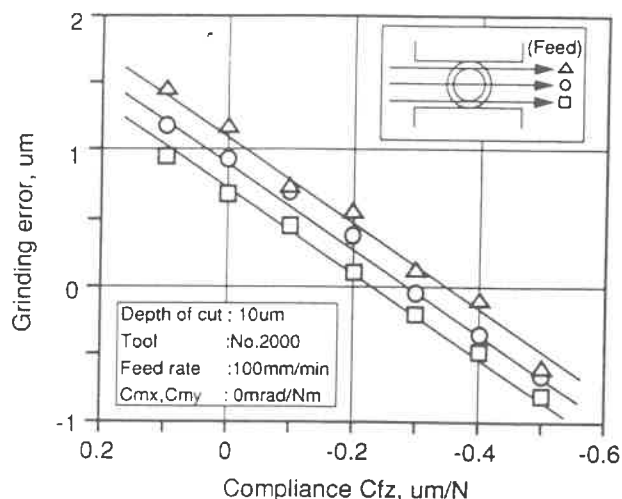


Fig.10 Compliance C_{fz} vs. grinding error

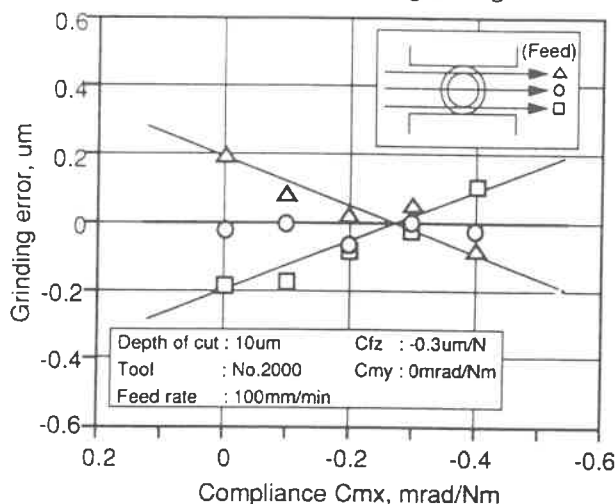


Fig.11 Compliance C_{mx} vs. grinding error

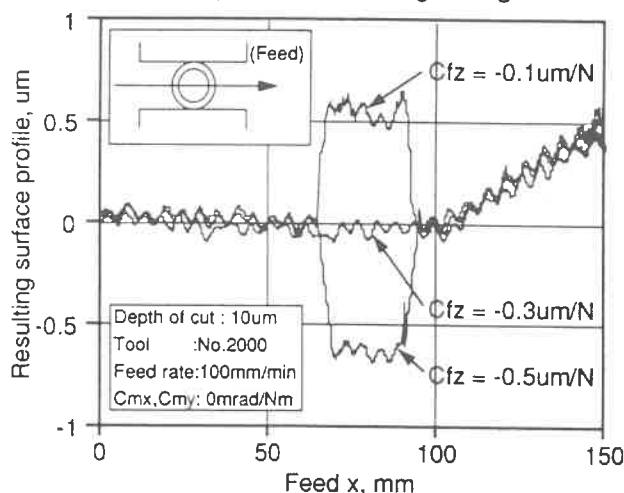


Fig.12 Resulting surface profile

Experimental Verification of a Grinding Chatter Model

Byron R. Knapp, Eric R. Marsh
The Machine Dynamics Research Lab
The Pennsylvania State University
University Park, PA

Introduction:

Grinding has long been used in the precision production environment where surface finish and geometric tolerances cannot be met with traditional cutting operations. With the introduction of 'near net shape' technology and the increasing demand for precision components, much larger demands will be placed on grinding performance. Once regarded as a finishing operation with low material removal rates, advances in abrasive technology and machine design will allow grinding to become the primary operation with large removal rates.

A major concern in any material removal process is the occurrence of chatter which causes relative motion between the tool and the workpiece. As a result of this motion, the workpiece surface will contain undesirable undulations and, in some cases, surface damage. In addition, the increased cutting forces associated with chatter lead to accelerated tool wear.

Many researchers have proposed different mathematical models in order to predict the occurrence of chatter. Several models of chatter have been developed for cutting processes (i.e., turning, milling, boring, etc.) including the two degree of freedom model proposed by Tlustý and Ismail [1]. However, there are several fundamental differences between cutting and grinding operations. For example, as opposed to a turning tool, grinding wheels contain many grains that are randomly spaced radially and around the periphery of the wheel. In addition, those grains can rub, plow, or cut the workpiece surface depending on wear, normal grinding force (F_n), and dress conditions. Furthermore, the development of lobes on both the wheel and the workpiece introduces a second feedback loop into the depth of cut which does not exist in the traditional cutting models. Because of these differences, cutting models cannot be applied to the grinding process.

Since the first grinding model by Hahn [2], several mathematical models have been devoted to predicting stability in grinding and the onset of chatter, but they lack published experimental verification. Recently, Thompson [3, 4] continued the development of his own doubly regenerative grinding chatter model but did not correlate it to experimental results. Verification of any mathematical model is necessary in order to determine if all the dominant parameters are correctly modeled. Once a grinding model is verified, its application in precision production grinding will increase the understanding of this

'black art' while improving cycle times and thus productivity. Unfortunately, experimental verification of grinding models is difficult to obtain and published verification is virtually non-existent. It is the aim of this work to compare Thompson's model to experimental results obtained from cylindrical plunge grinding tests.

The Model:

Two categories of chatter exist; forced and self-excited. Forced chatter is directly related to the wheel rotational frequency and can be caused by wheel imbalance, offset, or some other rotating imbalance in the machine. Vibrations independent of the wheel rotational frequency often correspond to a natural frequency of the grinding machine and are termed self-excited. Self-excited chatter will decay over time for a stable process and increase over time for an unstable one. Unstable, self-excited chatter, or regenerative chatter, causes lobing to occur on the wheel (wheel regenerative) or the workpiece (workpiece regenerative), or both. For example, *Figure 1* illustrates how regenerative chatter produces lobing in the workpiece.

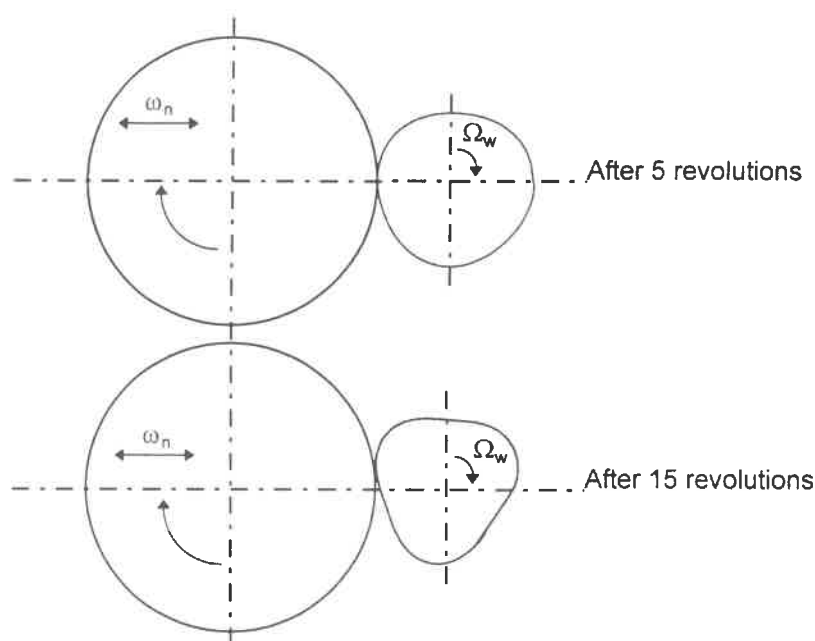


Figure 1. Workpiece regenerative chatter (after Salmon [5]).

The dominant mode of the machine is a horizontal vibration at $\omega_n = 18$ Hz and the workpiece rotational frequency is $\Omega_w = 6$ Hz (360 rpm) which eventually causes 3 lobes on the workpiece. Once the lobes form on the workpiece, they act as feedback into the system and can cause unstable vibration of the machine.

Thompson's model for regenerative chatter [3], shown in *Figure 2*, assumes a harmonic normal force, F_n , which may increase (decrease) over time according to the exponential growth (decay) index, α . Only the dominant mode of the machine is modeled by the effective m , c , and k , and Hertzian contact is neglected. Experimental verification of the model will help determine the relevance of these assumptions. Thompson solves the governing equations according to this model and proposes a triple iterative process for analyzing the growth of chatter.

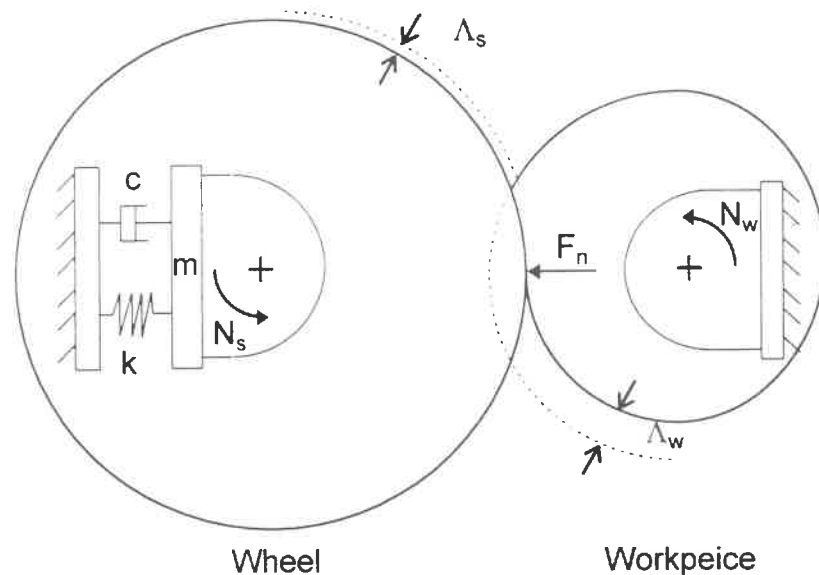


Figure 2. Grinding chatter model (after Thompson).

$$F_n = F_o e^{\alpha t} \sin \omega t$$

c	: machine damping	m	: equivalent machine mass
F_o	: initial force	t	: time
F_n	: normal grinding force	α	: exponential growth index
k	: machine stiffness	Λ_s	: wheel wear rate
N_s	: wheel rotation rate	Λ_w	: material removal rate
N_w	: workpiece rotation rate	ω	: chatter frequency

Verification:

The experiments were conducted as part of a joint investigation of several contributors to the Machine Dynamics Research Lab at Penn State. A Weldon 1632 cylindrical grinder with an Aluminum Oxide wheel was used to grind ANSI 8620 steel. The machine vibration was monitored during the process with several Kistler accelerometers. Measurement of the wheel spindle vibration was obtained with a Lion Precision capacitance gage. The dynamic characteristics of the grinder were obtained using experimental modal analysis, and the grinding process parameters were inferred from several grinding

tests. An exploration of the effects of several grinding parameters was made to determine the important variables in chatter and how they compare to the results of the triple iterative analysis proposed by Thompson.

Future Work:

The results of this study help relate theoretical predictions of grinding instability to actual conditions. Modifications of the model based upon these results will be made to insure accuracy. The model will then be applied in the production environment in order to supplement operator experience with experimentally proven theory, drastically reducing the time to manufacture and thus the cost to grind. In addition, the model will be used to monitor the stability of the grinding process which will decrease manufacturing time while increasing quality.

References:

- [1] **Tlusty, J. and Ismail, F.** "Basic Non-Linearity in Machining Chatter." *Annals of the CIRP*. Vol. 30, No. 1, 1981, p. 299.
- [2] **Hahn, R.S.** "On the Theory of Regenerative Chatter in Precision-Grinding Operations." *Trans. ASME*, Vol. 76, No. 1, 1954, p. 593.
- [3] **Thompson, R.A.** "On the Doubly Regenerative Stability of A Grinder: The Theory of Chatter Growth." *Journal of Engineering for Industry*, Vol. 108, May 1986, p. 75.
- [4] **Thompson, R.A.** "On the Doubly Regenerative Stability of A Grinder: The Mathematical Analysis of Chatter Growth." *Journal of Engineering for Industry*, Vol. 108, May 1986, p. 83.
- [5] **Salmon, S.C.** Modern Grinding Process Technology. McGraw-Hill, Inc.: New York. 1992. p. 152.

Keywords:

Grinding Chatter, Grinding Experimentation, Precision Cylindrical Grinding.

A Model for Self-Dressing of Resin Bound Grinding Wheels

Thrinath S. Kakumanu, Michele H. Miller
Michigan Technological University
Houghton, MI 49931

1 Introduction

The multiple cutters of a grinding wheel give grinding advantages over single point cutting, especially for machining hard work materials. While a single point tool will quickly dull and require re-sharpening, a grinding wheel distributes wear over many cutters and is capable of replenishing itself. Ideally, grinding wheel wear will occur fast enough to supply fresh cutting edges but slow enough to permit long wheel life and good machining accuracy. Understanding wheel wear and optimizing it is crucial to achieving good performance in the grinding of brittle materials.

For grinding brittle materials resin binders tend to perform better than metal binders. Their greater compliance puts more grits in contact with the workpiece and their lower hardness causes faster wear and more self-dressability. Both effects occur simultaneously, and it is unclear which effect is more important in terms of ductile grinding performance or if both contribute equally. In this paper we focus on the wheel wear and self-dressing of resin bound wheels.

2 Self-Dressing Process

Tang et al. observed two dominant wear mechanisms for resin bound wheels: grit pullout for blocky diamonds and grit fracture for friable diamonds [1]. Focussing on grit pullout as the primary mechanism, we develop a model to describe the self-dressing process and to predict the growth of the cut-

ting force. Self-dressing, a beneficial result of wheel wear, is a grit replenishment and wheel sharpening process. The cutting process causes sharp abrasive grit edges to wear. As a grit's wear plateau area increases, friction between the grit and workpiece increases causing an increase in the cutting force. When this cutting force exceeds the binder strength, the abrasive grit is pulled out of the binder matrix. The pullout of worn abrasive grits exposes fresh grits with sharp cutting edges.

Starting with a sharp dressed wheel, the overall grinding force will increase as the cutting grains dull. The force levels off and reaches an equilibrium when grits pull out and get replenished at a steady rate. The equilibrium force is significant in determining the surface finish of the workpiece and wear of the grinding wheel.

3 Model

Figure 1 outlines an algorithm that was developed to model the self-dressing phenomenon in plunge grinding. The algorithm initially estimates the number of cutting grits involved in the process. Then, it calculates grit forces over the entire wheel circumference as a function of wear plateau area (assumes one cutting edge per grit). It adds all the grit forces to get a total for one wheel revolution. Next, by comparing each grit force to a preset threshold force, it determines whether any grits would pull out of the binder. It then replaces pulled out grits with sharp grits that have a small wear plateau area.

The rate of self-dressing is influenced by: (1) the relationship between grinding infeed and number of active grits; (2) the relationship between grit sharpness and cutting force; (3) the relationship between binder hardness and the force required to pull a grit out of the binder.

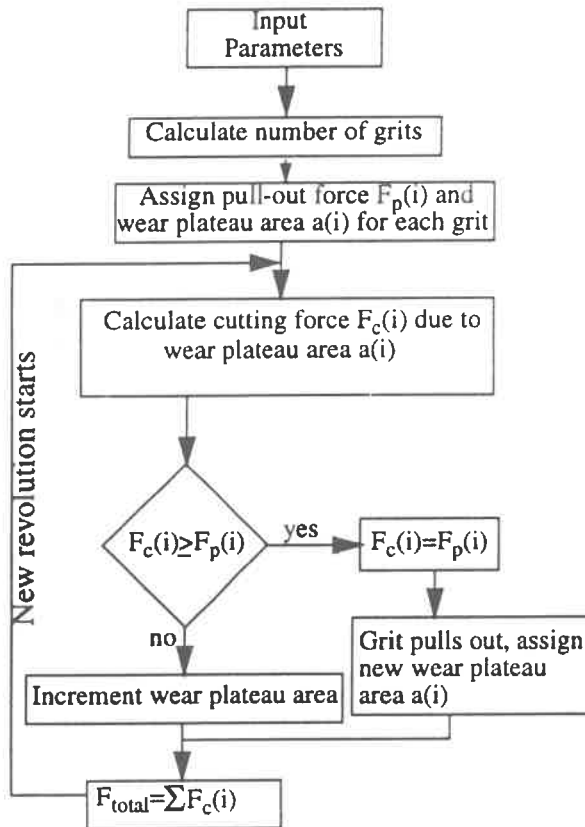


Figure 1: Flow chart for the self-dressing model

Effect of Infeed on Number of Cutting Grits Infeed rate affects the cutting force and therefore the rate of self-dressing. Increasing the infeed will increase the number of cutting grits and the force per grit. In our simplified model we neglect the change in grit force and consider only the impact on number of cutting grits. Based on the kinematics of plunge grinding, the number of cutting grits n_g changes with infeed according to [2]:

$$n_g = 0.52 \frac{w}{r_g^2} (f_v l)^{0.6} d^{0.4} \quad (1)$$

where,

d = infeed per revolution

w = width of grinding wheel

f_v = abrasive volume fraction

r_g = grit radius

l = wheel circumference

Effect of Wear Plateau Area on the Cutting Force According to Davis et al. [3], the normal grinding force has three components: (1) force on the flat wear plateau areas of the grits; (2) force on the curved portions of the grits; (3) force on the ploughing grits. For simplicity, we neglect the second and third components, and assume the normal force on one grit is related to work hardness as follows:

$$F_n = a H_w \quad (2)$$

where,

a = wear plateau area,

H_w = hardness of the workpiece

The tangential force is related to the normal force by:

$$F_t = \mu F_n \quad (3)$$

Therefore,

$$F_t = \mu H_w a \quad (4)$$

Wear plateau area a depends on the shape of the grit. Assuming a spherical grit, the wear plateau area is a circle. Defining the radius of the wear plateau as r_1 , then the area is:

$$a = \pi r_1^2 \quad (5)$$

Therefore the tangential cutting force is given by:

$$F_t = \pi \mu H_w r_1^2 \quad (6)$$

Here, μ depends on the type of abrasive and workpiece material used. Based on [4], a value of 0.44 was used in our model. The tangential cutting force then becomes,

$$F_t = 1.3816 H_w r_1^2 \quad (7)$$

Effect of Binder Hardness on Pull-out Force Pull-out force is defined as the force required to remove an abrasive grit from the binder matrix. In modeling the self-dressing phenomenon, measuring pull-out force for each abrasive grit is a significant task [5]. To derive this force, two parameters of the wheel were given importance--binder hardness and grit size. Assuming the abrasive grit as spherical and half embedded in the binder, the pull-out force was approximated as:

$$F_p = 2\pi r_g^2 Y_b \quad (8)$$

where Y_b is the yield strength of the binder.

4 Parametric Study

Results of the self-dressing model were studied to understand the effect of several different parameters. Figure 2 describes the effects of depth of cut, binder hardness, and workpiece hardness on the growth and stabilization of the cutting force. As infeed increases, the rate of force growth (slope) increases as well as the equilibrium force. The higher force is due to the larger number of cutting grits at high infeeds. Figure 2 also shows that harder work materials and harder binder materials lead to higher grinding forces.

5 Summary and Conclusion

The model presented here describes the self-dressing process and its effect on cutting force. It provides a qualitative picture of the effect of key grinding parameters. To improve the model, assumptions associated with wear rates and grit pull-out force will be further refined. Experiments will soon be conducted to assess the model's validity. In addition, more research is needed to relate the equilibrium force to the grinding performance. Given a desired equilibrium force, this model provides guidance in selecting infeed rate and binder hardness.

References

1. Tang, J., K. Syoji and D. Dornfield, "Diamond Wheel Wear Sensing with Acoustic Emission," *Proceedings of the ASPE Spring Topical Meeting*, Vol. 13, pp. 109-114, 1996.
2. Miller, M. H. and T. A. Dow, "Modeling the Wheel/Work Interface in Precision Grinding," *Proceedings of the ASPE Annual Meeting*, pp. 101-104, 1994.
3. Davis, C. E. and C. Rubenstein, "Reciprocating Plunge Surface Grinding with Diamond Abrasive Wheels," *Int. J. Mach. Tool Des. Res.*, Vol. 12, pp. 165-177, 1972.
4. Reinhart, *Proceedings of the International Industrial Diamond Conference*, Industrial Diamond Information Bureau, London, Vol. 2, p.19, 1967.
5. Peklenik, J. and H. Opitz, "Testing of grinding wheels," *Proc. 3rd Int'l MTDR Conference*, Pergamon Press, London, pp. 163-177, 1963.

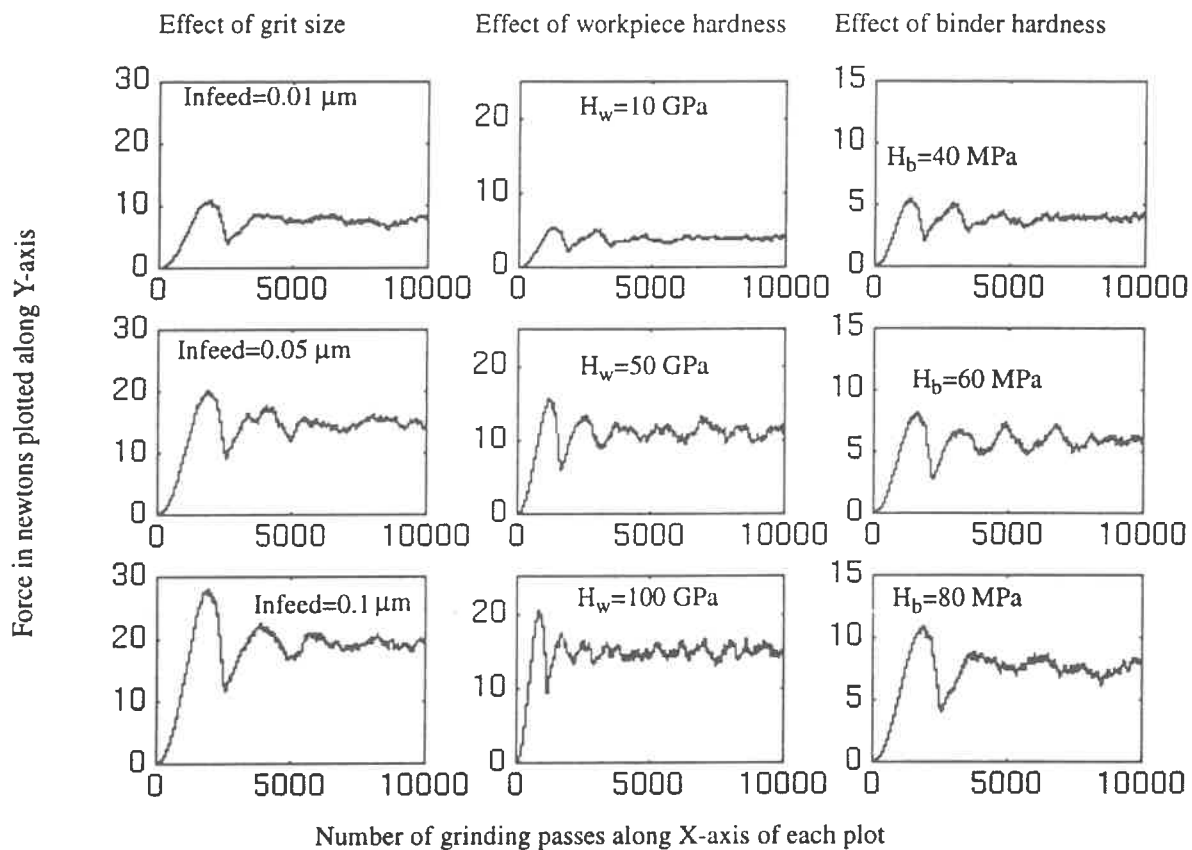


Figure 2: Effect of various parameters

THE EFFECT OF WORKPIECE MODULATION ON SURFACE QUALITY

Weimin Qu, Michele H. Miller
Department of ME-EM
Michigan Technological University
Houghton, MI 49931

1 Introduction

Ductile regime grinding is an enabling technology in the manufacture of ceramic components because it produces accurate shapes deterministically and without fracture. However, the process requires very small chip thicknesses which limit material removal rates (MRR) and consequently economic viability. Some researchers have taken advantage of ultrasonically-assisted machining to increase the MRR since the modulation can reduce the cutting forces. Moriwaki [1] achieved a seven times increase in critical depth of cut for single point machining of glass and credited the improvement to a reduction of frictional force on the tool rake face. However, the cutting speed was quite low--1.5m/s. With the much higher cutting speed of grinding and the short time of each abrasive/work interaction, it is uncertain whether high frequency vibration would have a similar effect on grinding performance.

In this paper we investigate the effect of workpiece modulation on grinding surface fracture. We also investigate the mechanisms by which modulation impacts ground surface quality.

2 Experiments

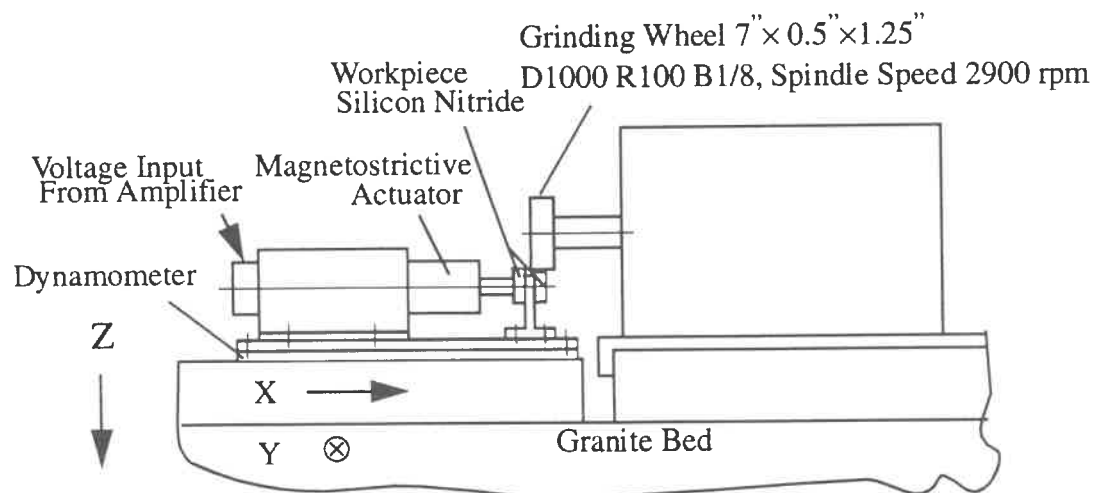


Figure 1: Schematic illustration of the surface grinder with the actuator

Experiments were conducted to assess the effect of modulation on the ground surface. Figure 1 shows a schematic illustration of the surface grinder with a magnetostrictive actuator. The work-

piece is silicon nitride and excited to vibrate in the crossfeed (or X) direction. Table 1 describes the experimental conditions.

Table 1: Experimental conditions

Spindle speed	2900 rpm
Cutting speed	29000 mm/s
Feedrate	13 mm/s
Depth of cut	0.015 mm
Workpiece	silicon nitride
Modulation frequency	2000 Hz
Modulation amplitude	2.2 μm
Grinding wheel	D1000 R100 B1/8

When the surfaces are ground, for each 15 μm depth of cut, the grinding wheel passes the workpiece four times. Figure 2 shows a surface ground with modulation, and Figure 3 shows a surface ground without modulation. Comparing the two figures, the surface ground with modulation has fewer scratches and less visible fracture.

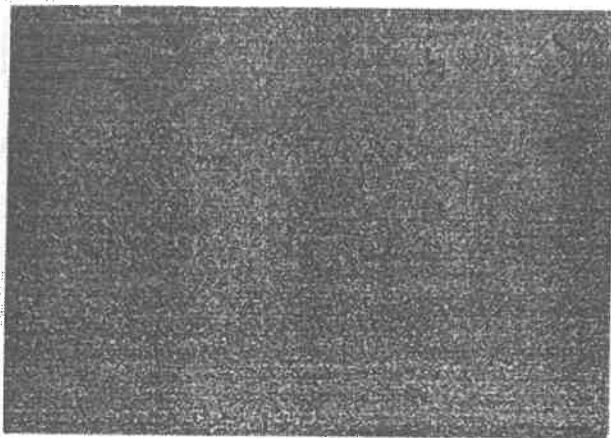


Figure 2: Surface texture with modulation(100X)

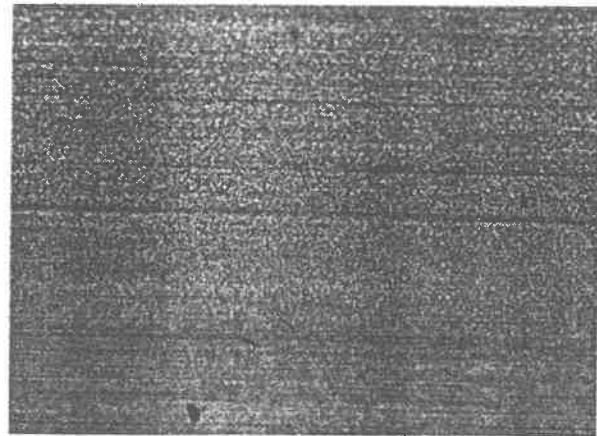


Figure 3: Surface texture without modulation(100X)

Surface profilometer traces confirm that the surface in Figure 2 has lower roughness than the surface in Figure 3. Table 2 summarizes the measurements.

Table 2: Surface roughness comparison

	Ra (μm)
with modulation	0.25
without modulation	0.39

3 Theory

Experiments demonstrate that high frequency modulation can improve the surface quality. To optimize the effect of modulation, we need to understand the mechanism behind its influence on

surface quality. Researchers have observed that high frequency modulation has the following effects:

- (a) reduction of the effective friction force between workpiece and cutting tool [2-6];
- (b) reduction of the effective yield strength of the workpiece [4,5];
- (c) self-dressing of the cutting grits [7];
- (d) reduction of surface temperature or better coolant penetration conditions [7,8].

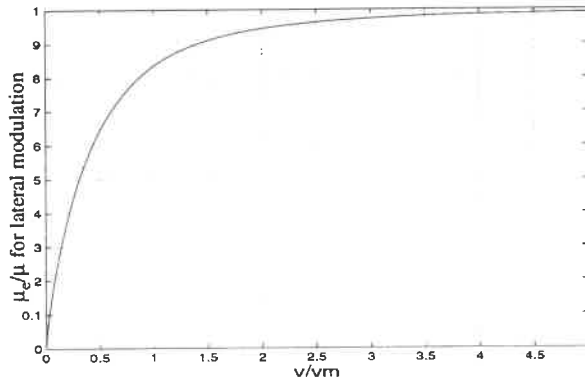


Figure 4: Reduction in friction coefficient as a function of velocity ratio

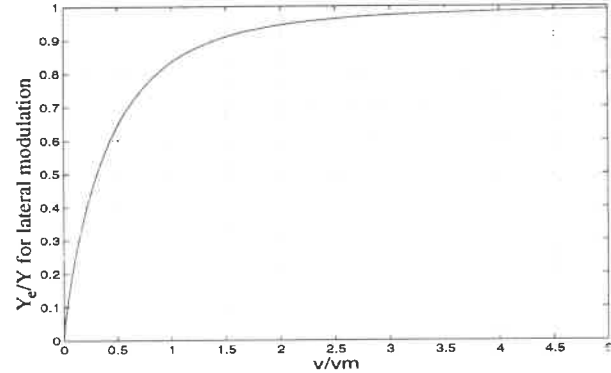


Figure 5: Reduction in flow stress as a function of velocity ratio

First, we consider the friction between the workpiece and a cutting grit under lateral modulation. According to [6], the vector of instantaneous friction force is constant in magnitude (independent of velocity magnitude) and in the direction opposite to the relative velocity. The lateral modulation thus redirects some of the frictional force to the lateral direction thereby reducing the component in the primary sliding direction. The primary direction component cycles between a nominal value (the value it would have without modulation) and some lower value. When averaged over time, the primary direction component is lower than the nominal value, and the lateral direction component is zero. The amount of reduction of the average or effective friction μ_e with respect to the nominal μ depends on the ratio of the cutting velocity v and the modulation velocity v_m . Figure 4 shows the variation in friction coefficient ratio μ_e/μ with speed ratio v/v_m . According to Figure 4, modulation has little effect on friction when the sliding velocity is much greater than the modulation velocity. For the experimental conditions described in Section 2, in which the cutting speed was approximately 1000 times the modulation velocity ($v = 29$ m/s and $v_m = 0.0276$ m/s), the reduction in friction is only 0.02%. Even with a higher modulation frequency of 20 kHz and larger modulation amplitude of 10 μ m, the friction reduction would only be 0.07%.

Second, we consider the lateral modulation effect on the reduction of the effective yield strength. The cutting force is approximately equal to the work material's flow stress multiplied by the tool area projected into a plane perpendicular to the cutting direction. In the presence of lateral modulation, the instantaneous cutting velocity changes in direction and magnitude. However, for a work material whose flow stress is independent of cutting velocity and for a cutting grit whose projected area is independent of cutting direction, the instantaneous cutting force changes in direction but not magnitude. Similar to the friction effect described above, some of the cutting force is redirected into the lateral direction. The average force in the primary cutting direction is,

therefore, lower than the nominal force. In other words, the material appears to have a lower flow stress. The reduction in this effective flow stress Y_e with respect to the nominal Y depends on v/v_m as shown in Figure 5. Considering the velocity ratio in our experiments (or even one achieved by ultrasonic modulation), Figure 5 indicates that the flow stress effect is negligible.

Regarding the third possible explanation for improved ground surface quality, Colwell found that 10 kHz vibrations inhibited the build up of debris on the grinding wheel [7]. In the honing of steels, Markov found that 18 kHz vibrations permitted better coolant penetration [8]. Further experiments are needed to test the validity of these explanations.

4 Summary

From an experimental study of the modulation grinding of silicon nitride we found that grinding with modulation reduces surface roughness and fracture. Analysis shows that two of the possible explanations for improved surface quality--reduced effective friction and reduced effective flow stress--are invalid. Additional experiments that focus on the wheel condition and workpiece temperature will be conducted.

References

- [1] Moriwaki, T., E. Shamoto and K. Inoue, "Ultraprecision Ductile Cutting of Glass by Applying Ultrasonic Vibration," *Annals of the CIRP*, Vol. 41, No. 1, pp. 141-144, 1992.
- [2] Astashev, V. K., "Effect of Ultrasonic Vibration of a Single-Point Tool on the Process of Cutting," *Journal of Machinery Manufacture and Reliability*, No. 3, pp. 65-70, 1992.
- [3] Moriwaki, T. and E. Shamoto, "Ultrasonic Elliptical Vibration Cutting," *Annals of the CIRP*, Vol. 44, No. 1, pp. 31-34, 1995.
- [4] Cornier, A. "Turning Assisted by Non Conventional Process," *Bulletin Du Cercle D'etudes Des Msetaux*, Vol. 16, No. 4, pp. 5.1-5.16, 1992.
- [5] Markov, A. I., *Ultrasonic Machining of Intractable Materials*, Ilife Books, London, 1966.
- [6] Nazarenko, A. F. and E. G. Topilin, "Study of vibrations effect on friction," *Journal of Friction and Wear*, Vol. 15, No. 4, pp. 36-39, 1994.
- [7] Colwell, L. V. "The effects of high-frequency vibrations in grinding," *Transactions of the ASME*, pp. 837-846, May, 1956.
- [8] Markov, A. I. and P. A. Ermak, "Ultrasonic diamond honing of steels," *Vestnik Mashinostroeniya*, Vol. 62, No. 4, pp. 52-55, 1982.

Grinding With Directionally Aligned SiC Whisker Wheel

– Loading-free Grinding –

Katsumi YAMAGUCHI*, Iwao HORAGUCHI** and Yutaka SATO***

* Department of Mechanical Engineering, Nagoya University,
Furo-cho, Chikusa-ku, Nagoya 464, Japan

** Toyota College of Technology, Eisei-cho 2-1,
Toyota-shi, Aichi-ken 471, Japan

*** Aichi Industrial Research Institute,
Nishishinwari, Hitotsugi-cho, Kariya-shi, Aichi-ken 448, Japan

1. Introduction

In a previous work [1], the authors developed a SiC whisker grinding wheel with its whiskers aligned normally to the grinding surface. Using this new type of grinding wheel, a surface finish of 16 nmR_{max} (1.5 nmR_a) roughness was successfully produced in hardened die steel. The SiC whisker is characterized by its small size and high aspect ratio, for example, a diameter of 1 μm with a length of 40~50 μm . This high aspect ratio helps the whisker to bond strongly and to be difficult to tear out from the grinding wheel surface. These unique characteristics may lead to longer durability and finer grinding over the conventional grain-bonded grinding wheel.

An anticipated weak point of the SiC whisker grinding wheel is that it is apt to load because the whiskers cannot be renewed in the grinding process as the abrasive grains can in conventional grinding wheels.

In the present research, a simple method to prevent loading is proposed, named "Loading-free grinding". In this method, a grinding fluid mixed with a small amount of very fine abrasive grains is used. The authors investigated the effects of the size and content of grains added to the grinding fluid on grinding characteristics. The ground surface and the wheel surface were observed under scanning electron microscope to verify the action of

the grains in the grinding fluid.

2 Experimental Detail

2.1. Loading-free grinding

In this experiment, loading-free grinding is proposed, by which loading may be completely prevented. The principle of the loading-free grinding is illustrated in Fig. 1, where it is supposed that the grinding fluid brings the grains into the contact surfaces, and the grains strike the adhered chips and the binder. By this motion, the grains may tear out the grinding chips adhered to whiskers and abrade the binder in the process, hence improving grinding efficiency.

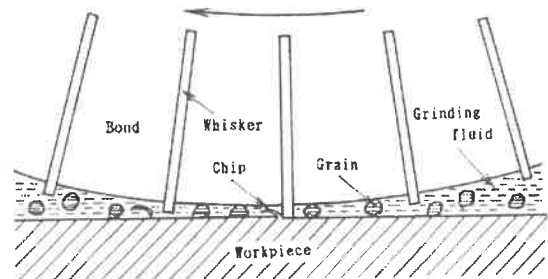


Fig. 1 Principle of loading-free grinding

2.2. Grinding wheel, fluid and workpiece

The grinding wheel used in the experiments is composed of SiC whisker as the abrasive and phenol resin as the binder. The whiskers were mixed with phenol at a ratio of 14% in volume. The whiskers were aligned

unidirectionally using the method reported in the previous paper [1]. Then the mixture was baked at a controlled temperature at a maximum 180°C for 30 hours. The prepared material was then cut into smaller blocks and fixed by epoxy binder on a cup-shape holder to align the whiskers in the direction normal to the grinding surface as demonstrated in Fig. 2. Next, parallel grooves were provided on the grinding wheel surface to facilitate the removal of grinding chips from the grinding surface.

The grinding fluid was soluble type (JIS W2-1), in which alumina (WA) grains of 0.4 ~ 1.2 μm in diameter are mixed to prevent the wheel from loading.

The workpiece for the grinding test was hardened die steel SKD11 of Rockwell hardness HRC60, with an initial surface roughness of about 1.5 μmRa .

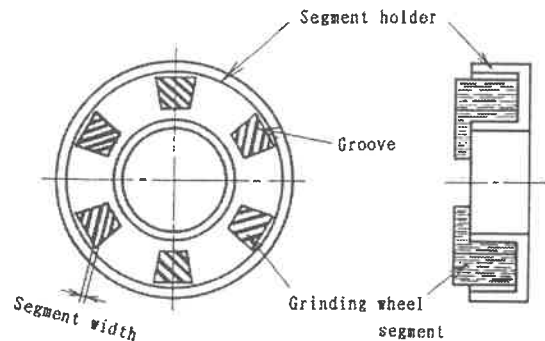


Fig. 2 Grinding wheel.

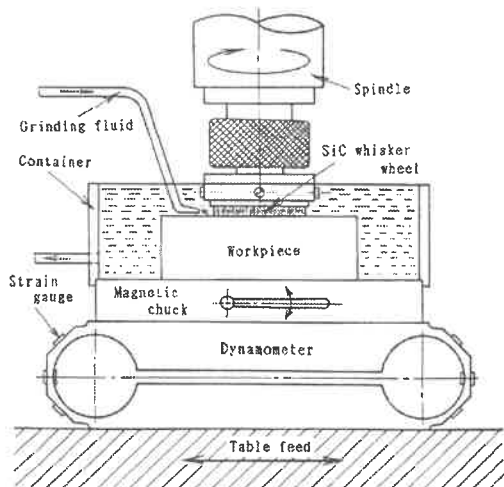


Fig. 3 Experimental set-up.

2.3. Experimental apparatus

The apparatus used for grinding experiments was a CNC vertical milling machine as illustrated in Fig. 3. The cup-shape grinding wheel was fixed on the machine spindle. The grinding was performed in the grinding fluid. The grinding test was done under a constant grinding load. Fig. 4 shows a SEM micrograph of the surface of the grinding wheel after grinding. The arrow indicates a representative tip of an aligned whisker, while the boxed X shows some whiskers clogged with chips.

To measure the ground volume from the workpiece, micro-Vickers indentations were marked on the work surface. The ground volume was measured from the changes in diameter of Vickers marks. As for the wear of grinding wheel, the Vickers indentations were marked on small pieces of brass, which were inserted in the segments. The experimental conditions are summarized in Table 1.

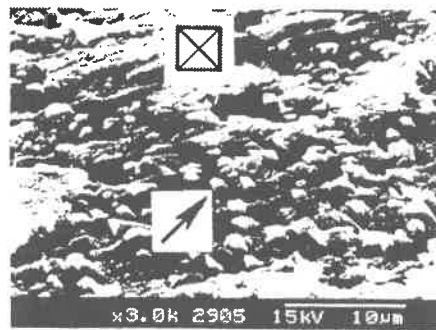


Fig. 4 SEM micrograph of grinding wheel structure

Table 1 Experimental conditions

Experimental apparatus	CNC milling machine
Grinding conditions	Wheel speed: 50 m/min
	Load: 400 N
	Feed: 35-42 mm/min
Grinding wheel	Wheel diameter: 50 mm
	Whisker content: 14 vol%
	Segment width: 1.5-8 mm
Specimen	Die steel SKD 11 (HRC60)
Abrasive grain	Material: Al_2O_3 (WA)
	Grain size: 0.4-1.2 μm
	Concentration in fluid: 0-0.5 wt%
Grinding fluid	Soluble type (JIS W 2-1)

3 Characteristics of Loading-Free Grinding

3.1. Effect of grain size

The results of surface finishes are compared for different sizes of WA grains in Fig. 5. The initial surface roughness was about $1.5 \mu\text{mR}_{\text{max}}$ and the concentration of WA grains in the fluid was 0.25%wt. The surface roughness R_{max} is plotted against the traverse cycles of the grinding wheel. From this figure, it is noticeable that, with grains of $0.6 \mu\text{m}$ size or $1.2 \mu\text{m}$, a surface finish can be obtained faster than with the finer grains of $0.4 \mu\text{m}$ in size. This means that the grains equal to or just smaller than the whisker diameter are the best in improving the anti-loading or the grinding efficiency.

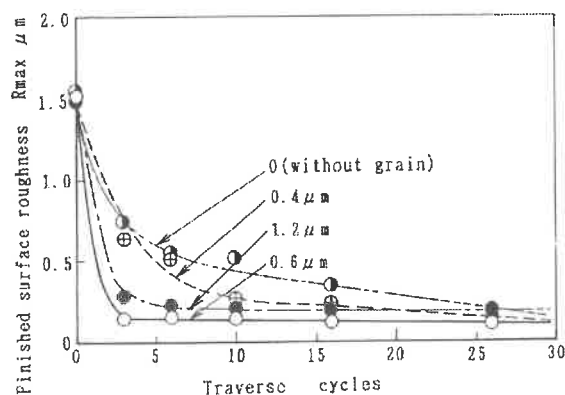


Fig. 5 Effect of WA grain size
(Concentration: 0.25wt%,
Segment width: 1.5mm)

3.2. Effect of grain concentration

The effect of grain concentration on anti-loading was investigated using alumina grains of $0.6 \mu\text{m}$ size. The material removal rates from the workpiece and wear of grinding wheel were shown in Fig. 6. The removal rate from the workpiece was found to increase, first, sharply with the increase in grain concentration, and then to saturate at a concentration around 0.25%wt. As for the grinding wheel, the wear rate always increases

at an extremely slow rate. Hence, a WA grain concentration of 0.25%wt was found to be the most suitable and sufficient ratio for anti-loading.

The scanning electron microscopy of the wheel surface, shown in Fig. 7, indicates that a low concentration of grains can remove the grinding chips, which allows the whiskers to protrude on the wheel surface so as to improve the anti-loading.

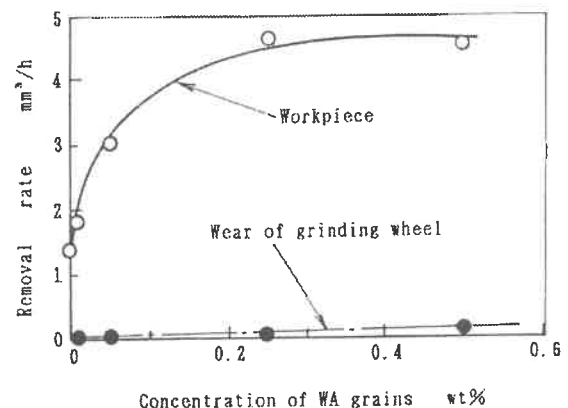


Fig. 6 Removal rates with different grain concentrations. (Grain size: $0.6 \mu\text{m}$)

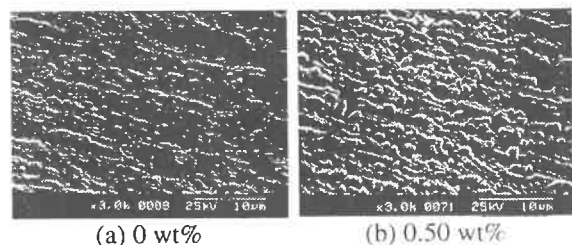


Fig. 7 Effect of abrasive grain concentrations on loading.

3.3. Grinding ratio

In Fig. 8, the ground volume of the workpiece and the wear of the grinding wheel are compared in two cases with and without anti-loading grains. The grinding rate is increased 10 times with the addition of grain to the grinding fluid, and the wear of the grinding wheel is increased. However, the grinding ratio is increased from 1000 to 6000 for two cases of grinding without and with anti-loading grains, respectively.

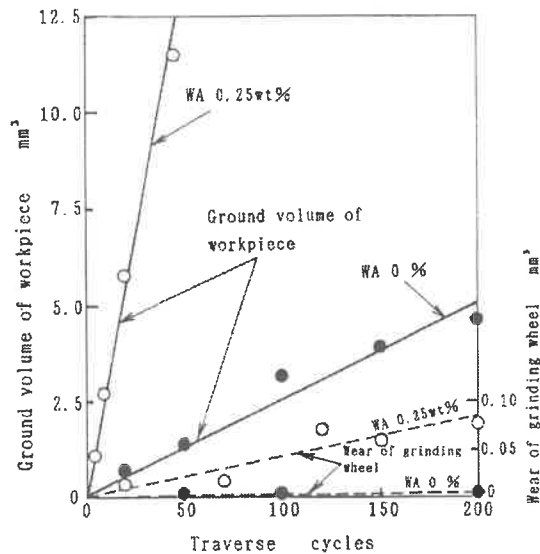


Fig.8 Grinding characteristics
(Grain size: $0.6\mu\text{m}$)

3.4. Ground surface observation

Fig. 9 shows AFM photographs of the ground surfaces for two cases; a) the surface obtained without using anti-loading grains,

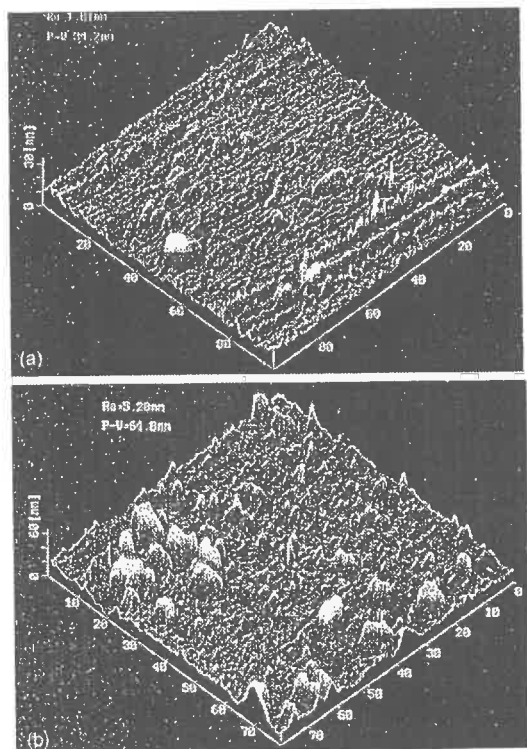


Fig. 9 AFM micrographs of ground surfaces.
(a) Without anti-loading
(b) With anti-loading (WA 0.25 wt%)

where a roughness of 16 nmRmax (1.5 nmRa) could be accomplished; and b) the surface obtained by applying anti-loading grains. The latter surface has a roughness of maximum 100 nmRmax (10 nmRa).

4 Conclusions

The results of this work are as follows:

- 1) A loading-free grinding is proposed in this paper. The loading problem in grinding with SiC whisker grinding wheel has been successfully resolved by this method.
- 2) The addition of a small amount of fine abrasive grains to the grinding fluid works as an anti-loading agent that can remarkably increase grinding efficiency.
- 3) The optimum grinding efficiency can be obtained by using anti-loading grains just smaller than the whisker diameter.
- 4) The concentration of 0.25%wt of anti-loading grains is quite sufficient. A higher concentration results in greater grinding wheel wear without improving the grinding efficiency.

References

- [1] K. Yamaguchi and I. Horaguchi. Development of directionally aligned SiC whisker wheel. Precision Engineering, vol.17, no.1 (1995) 5-9.

AN APPLICATION OF 40kHz ULTRASONIC VIBRATION TO SHEAR-MODE GRINDING OF SILICON WAFERS

Y. Yoshida *, K. Imai**, H. Hashimoto**, Z. N. Sun[○], Y. S. Lu^{○○}

*Tokyo Diamond Tool, ** Kanagawa Institute of Technology

[○]Harbin Institute of Technology, ^{○○}Northeastern University

1030 Shimo-ogino Atsugi-shi, 243-02 JAPAN TEL0462-41-1211 FAX0463-93-5440

Introduction

Previous papers ^{1), 2)} reported shear-mode grinding of various brittle materials using #3000 grit metal bond diamond straight wheels. In those experiments, it was shown that: (1) After a number of passes, the normal grinding force reached an approximate steady state: (2) During grinding, the wheel surface roughness and cutting performance were invariant, and wheel truing and dressing were not carried out.

From these experimental results, the shear-mode grinding force criterion was defined as the steady, invariant grinding force that is a repeatable deterministic function of shear-mode grinding conditions including materials and grinding wheels.

The criterion can usefully be applied to design a high precision shear-mode grinding machine by determining its compliance, and enables shear-mode grinding of various brittle materials without any in-process dressing.

The cup wheels can manufacture large surfaces efficiently and the surface accuracy is stable. However, the criterion isn't sufficiently defined yet using cup wheels. Because the contact area between the wheel and work increases, there are many problems such as the supply of coolant and the occurrence of grinding heat in the cup wheel. The ultrasonic vibration helps the supply of coolant between the wheel and work. Usually ultrasonic vibration was applied to high efficiency machining of hard and brittle materials with brittle-mode.

In this paper, the application of 40kHz ultrasonic vibration to shear-mode grinding is described. However, shear-mode grinding was expected to be possible in using fine grit and ultrasonic vibration of small amplitude. And it is performed in order to realize shear-mode grinding of sixteen inch silicon wafers, which are expected to be manufactured in the twenty-first century.

Experimental method

In experiments, metal bond cup wheels were used on a reconstructed surface grinding machine (Okamoto PSG-52DX) with a Professional Instrument's "Block-Head" type 4R. The ultrasonic spindle was installed under the air spindle. The frequency of the ultrasonic vibration is 40 kHz and the amplitude is 2 μ m. Batches of 10 $\times$ 25mm silicon wafers (100 surface) were ground with ("ultrasonic vibration grinding") and without ultrasonic vibration ("conventional grinding"), with various diamond grits, #170~#3000. The wheel rotation was changed to 1000 and 2000 rpm, the table speed was

changed to 50~150mm/min and the depth of cut was changed 0.5~4 μ m/pass.

The schematic illustration of the grinding machine is shown in Figure 1 and the grinding conditions are shown in table 1.

Definition of the grinding force

The grinding force contains the high frequency ingredient of 40 kHz. Therefore, to examine grinding force in this paper, we joined an amplifier to a dynamometer and examined the value which is output by the digital oscilloscope and the pen recorder in the comparison. As a result, the

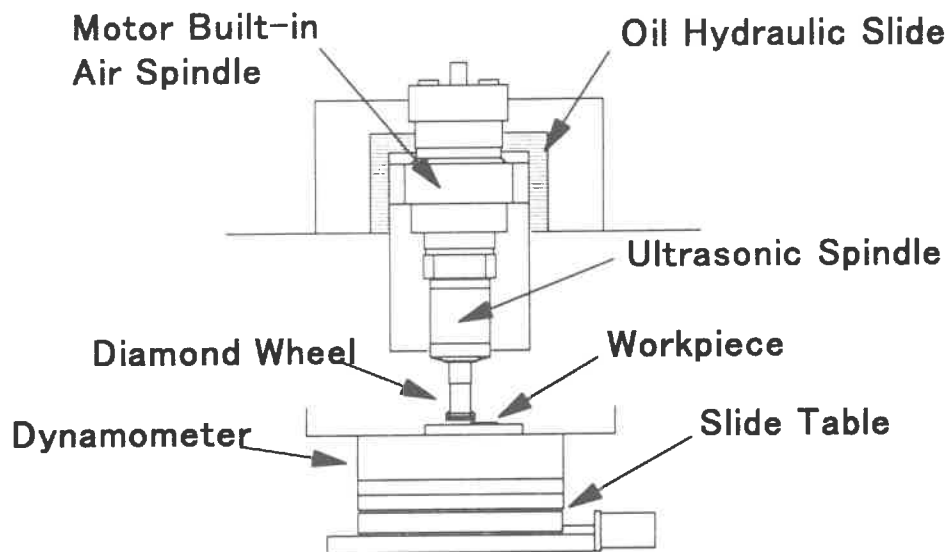


Figure 1 Schematic Illustration of The Grinding Machine

Table 1 Experimental Equipment and Grinding Conditions

Grinding machine	PSG-520DX (Okamoto Machine Tool)
Air spindle	BLOCK-HEAD Type 4R (Professional Instruments)
Slide table	Twinkler 5000 (THK)
Dynamometer / Charge Amplifier	TYPE 9257A / TYPE 5007 (KISTLER)
Work piece	Silicon 10×25×2mm (100 surface)
Diamond wheels φ 42-1w (cup type)	SD#170 P 75 M SD#1000 P 75 M SD#3000 P 75 M
Ultrasonic Spindle	
Frequency	40kHz
Amplitude (wheel surface)	2 μ m
Grinding Conditions	
Wheel Rotation (Peripheral Speed)	1000, 2000 rpm (2.2, 4.4 m/s)
Table Speed	50, 100, 150 mm/min
Depth of Cut	0.5, 1, 2, 4 μ m/pass

corrugation which corresponds to 40 kHz and the high frequency ingredient of about 4 kHz were observed by the oscilloscope. In the pen recorder, it was found that the value which is equivalent to the integration value which contains these high frequency components was put out. There are a few changes, though we found that grinding force according to depth of cut was measured using the dynamometer and the pen recorder. Therefore, in this paper, we defined the value which was registered by the pen recorder as being the grinding force value.

Experimental results

Changes in normal grinding forces are shown in Figure 2, grinding conditions are as follows: Wheel rotation: 2000rpm (2.2m/s), table speed: 100mm/min, and depth of cut: $1 \mu\text{m/pass}$. In the figure, white marks show conventional grinding, and black marks show grinding with ultrasonic vibration. Brittle-mode surface was produced by #170 and #1000 grit wheels with ultrasonic vibration. Then the normal grinding forces were lower than normal grinding force by conventional grinding. But a #3000 wheel produced a high normal grinding force by grinding with ultrasonic vibration. The normal

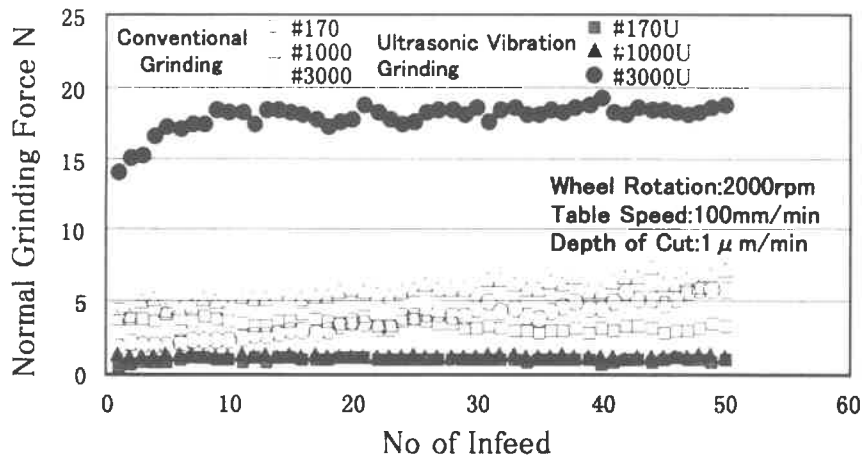


Figure 2 Comparison of normal grinding forces of various grit size

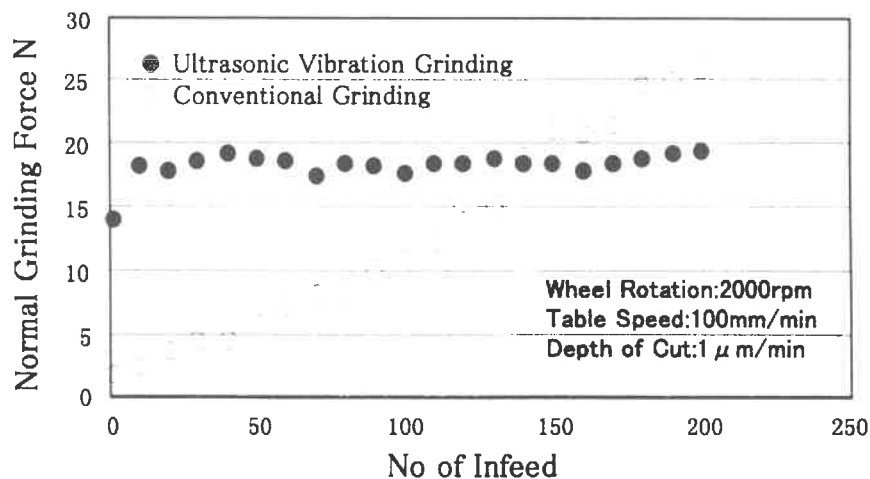


Figure 3 Comparison of normal grinding forces with a #3000 grit wheel

grinding forces by conventional grinding with #1000 and #3000 grit wheels increased with passes, but they were stable in grinding with ultrasonic vibration. In the shear-mode grinding, the condition of the wheel surface changed and the normal grinding force rose with an increase in the number of passes.

However, because there is little change of the wheel surface in grinding with ultrasonic vibration, we think that the normal grinding force will change very little. With a #3000 grit wheel, in case of ultrasonic vibration grinding, the tangential grinding force was stable and smaller than with conventional grinding. The value was about 1/5 in comparison with conventional grinding. This fact is extremely favorable in shear-mode grinding with a cup wheel.

Comparison of change of the normal grinding force with and without ultrasonic vibration under 200 passes is shown in Figure 3. The normal grinding force of conventional grinding increased with passes and exceeded the value, produced by ultrasonic vibration grinding. On the other hand, the normal grinding force of grinding with ultrasonic vibration showed comparatively high value, but it was quite stable.

The micrographs of the ground surfaces with various grinding conditions and various grit sizes are shown in Figure 4. Ground surfaces by both grinding methods in #170 grit wheel were approximately brittle-mode. With a #1000 grit wheel, the ground surfaces by conventional grinding were approximately shear-mode. On the other hand the surfaces by ultrasonic vibration grinding were intermingled with shear-mode and brittle-mode. We consider that the ground surface of conventional grinding was affected by the grinding heat, but the ground surface of ultrasonic vibration grinding was little affected by the grinding heat according to the observation results of the micrographs. When using #3000 grit wheel, we could produce shear-mode surfaces in any condition.

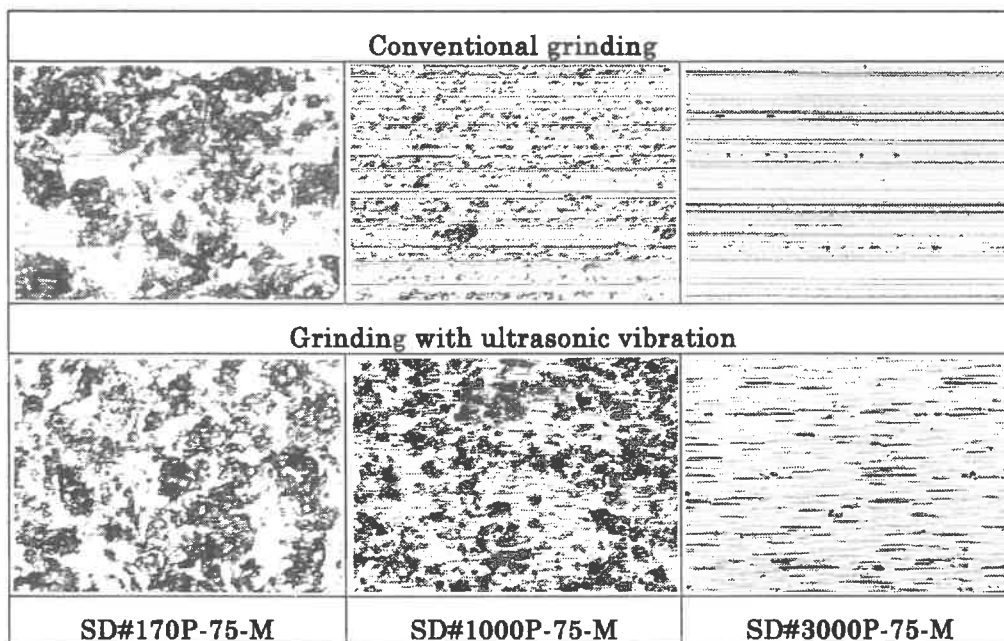


Figure 4 Micrographs of ground surfaces with various grit size and both grinding methods

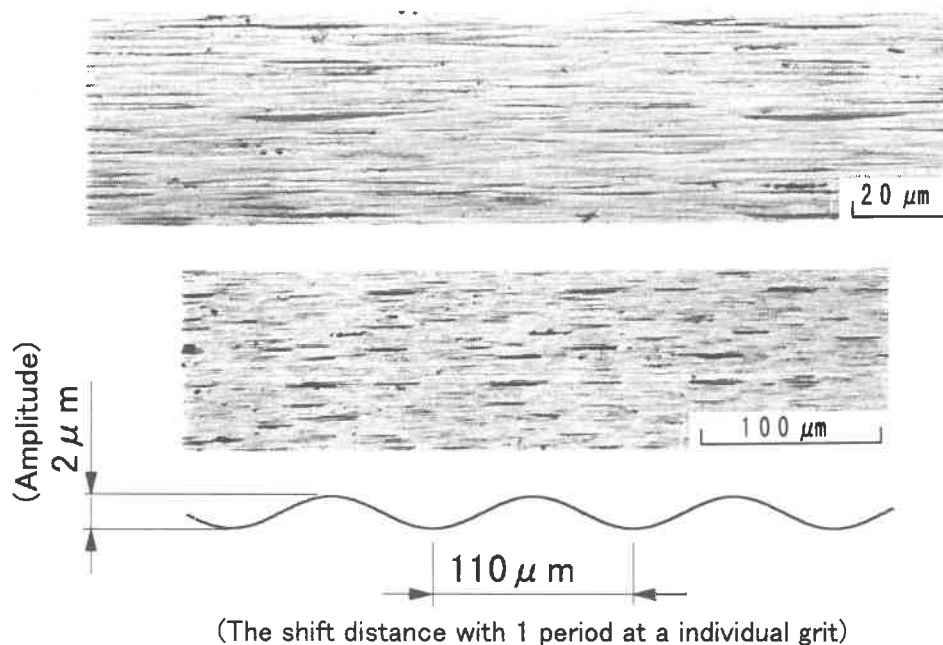


Figure 5 Micrographs of grinding streaks
and the track of theoretically calculated individual grits

The grinding streaks which were shown in the micrographs agree with the track of theoretically calculated individual grit. We think the streak was produced by rough or worn-out grit. It is necessary to produce finer streaks; therefore, the magnitude of the individual grit machining unit needs to be small.

Conclusions

We found the following in silicon wafer grinding with ultrasonic vibration:

- ① Under various conditions, with and without ultrasonic vibration, shear-mode surface can be produced with #3000 grit diamond cup wheels.
- ② The grinding force was kept dramatically stable during many repeated passes in grinding with ultrasonic vibration.
- ③ It is possible to produce a smoother surface by using a finer diamond grit with ultrasonic vibration grinding.

These experimental results show great potential for using ultrasonic vibration in manufacturing next generation sixteen-inch silicon wafers.

Reference

- 1) H. Hashimoto, K. Imai, 10th Proc. ASPE, (1995)171
- 2) H. Hashimoto, J. Takeda, K. Imai, K. Blaedel, Int. JSPE, 27,2 (1993)95.

SURFACE AND SUBSURFACE STRUCTURE OF ALUMINA AFTER DUCTILE-MODE GRINDING

I. Zarudi and L. C. Zhang¹

Department of Mechanical and Mechatronic Engineering, The University of Sydney, NSW
2006, Australia

1. Introduction

Alumina is known as a material of low density with high temperature strength and high resistance to wear and corrosion. These properties make it attractive to industry for producing electronic and optical components that require high quality of surface and damage-free subsurface. Fortunately, the recent application of ultra-precision grinding [1-2] has made it possible to produce surfaces with high integrity. However, the formation mechanism of the subsurface structure has not been well understood yet and its correlation with grinding conditions has not been fully established. As pointed by the research group in the University of Sydney, the structure of subsurface plays a crucial role in affecting the properties of a ground component and microcracks may exist in the subsurface even if the ground surface is absolutely microcrack-free.

The purpose of this work is to gain an understanding of the formation mechanism of the surface and subsurface structures of alumina in ultra-precision grinding.

2. Method

The materials used for study were alumina with average grain diameters of 1 μm and 25 μm . The grinding experiments were conducted on an ultra-precision surface grinder, Minini Junior CF CNC M286. A diamond grinding wheel of SD4000L75BPF(diameter 305 mm, peripheral speed 27 m/s) was used. The depth of cut was fixed at 100 nm but the table speed was varied from 0.02 m/min to 1 m/min. A water based coolant Syntilo 3 (99 % water, 1 % mineral oil) was used.

Topography of the ground surfaces was explored by means of High Resolution Scanning Microscope (HRSEM) JSM-6000F. The surface roughness was measured by Atomic Force Microscope (AFM) and the subsurface structure of specimens was studied on Transmission Electron Microscope (TEM) EM 430. Detailed procedure can be found in Ref. [3].

3. Results and Discussion

3.1 Overall features of ground surfaces

The surface topography of ground alumina is shown in **Fig.1**. Mirror surfaces were obtained under all the table speeds used. Plastic grooves were clearly seen on the ground surface, suggesting the ductile mode of material removal. However, some pits were also observed. The

¹ Author for correspondence

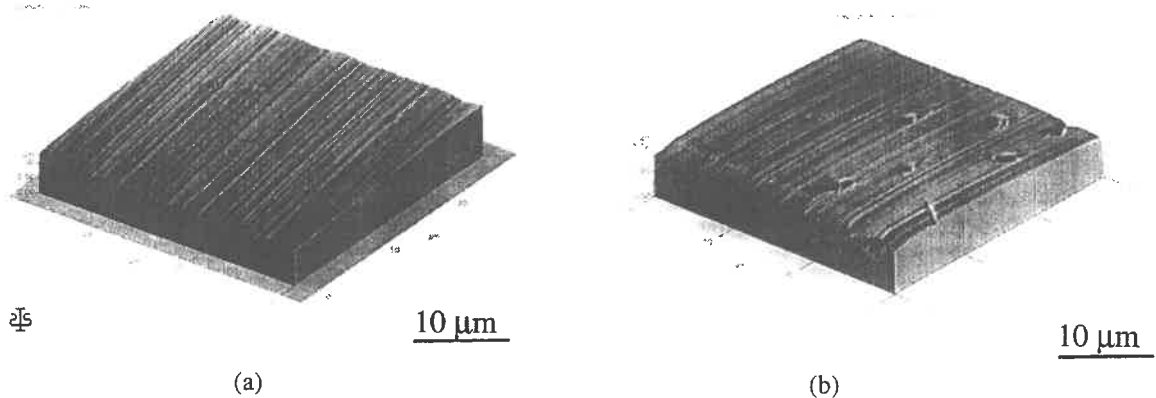


Fig.1 Surface topography after ultra-precision grinding (Table speed 0.02 m/min): (a) The 1 μm-grained alumina; (b) The 25 μm-grained alumina.

variation of Rms roughness vs table speed for both types of alumina is shown in **Fig.2**. It is clear that Rms roughness increased with the increase of table speed. For the 1 μm-grained alumina Rms roughness increased from 30 nm to 50 nm with the increase of table speed from 0.02 m/min to 1 m/min. For the 25 μm-grained alumina the effect was more pronounced and Rms roughness varied from 33 nm to 90 nm. Clearly, Rms roughness was affected by pits on the surface. The surface covered by pits was determined by the technique of image analysis, as shown in **Fig.3**. It is obvious that the surface covered by pits decreased significantly with the decrease of table speed, i.e., with the decrease of the uncut chip thickness. Furthermore, the surface covered by pits in the 1 μm-grained alumina (1.5 % for table speed 0.02 m/min) is less compared with those in the 25 μm-grained alumina (5 % for the same table speed).

To study the nature of pits we explored their topography by HRSEM. As shown in **Fig.4**, the central part of a pit resembles the cross-section of pores that were formed during specimen processing. The peripheral part of the pit possesses the characteristics of fracture.

The fracture around the pit can be explained by the effect of pore on the grinding process. A pore in the material can act as a stress raiser and promote microcracking near its

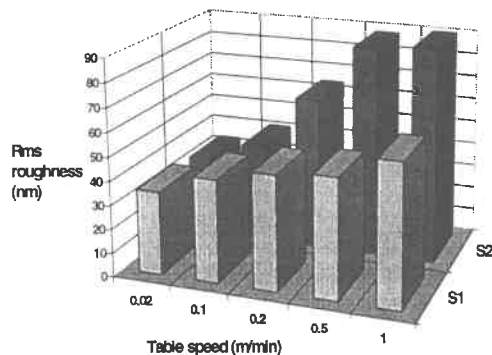


Fig.2 Effect of table speed on Rms roughness in: (S1) The 1 μm-grained alumina; (S2) The 25 μm-grained alumina.

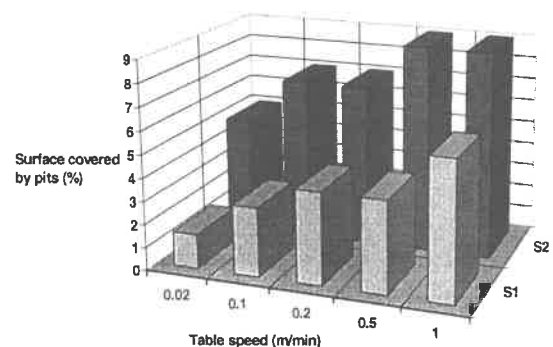


Fig.3 Effect of table speed on surface covered by pits in: (S1) The 1 μm-grained alumina; (S2) The 25 μm-grained alumina.

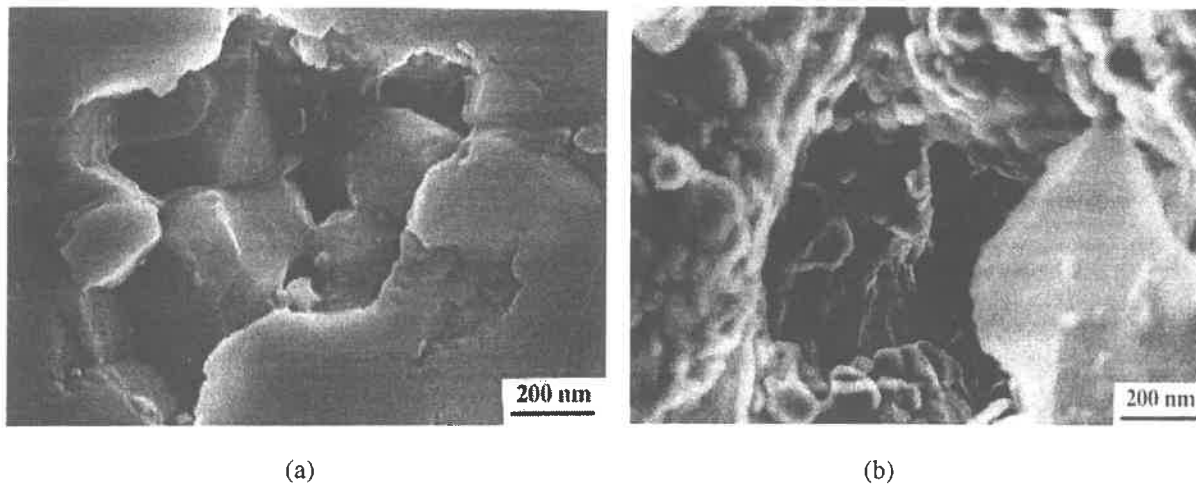


Fig.4 Topography of a pit in: (a) The 1 μm -grained alumina; (b) The 25 μm -grained alumina.

edge during the abrasive cutting of grinding wheel. It is worth to note that the fractured surface area of the 25 μm -grained polished alumina was 18 %, this is much higher than that of the ground 25 μm -grained alumina (9 %). This means that ultra-precision grinding can generate better surface quality compared with ordinary polishing.

3.2. Subsurface structure

Before ultra-precision grinding, the damaged subsurface layer induced by specimen preparation (conventional grinding) was determined so that this initially damaged zone could be removed completely during the subsequent ductile-mode grinding. Thus in all the cases below the subsurface structures were purely caused by grinding.

Figure 5 shows that ductile-mode grinding created a layer with an extremely high density of dislocations, which was immediately beneath the ground surface. This layer spread to the depth of 100 nm to 150 nm under the surface depending on table speed. Following this, a zone with a much lower density of dislocations appeared. Dislocations could be

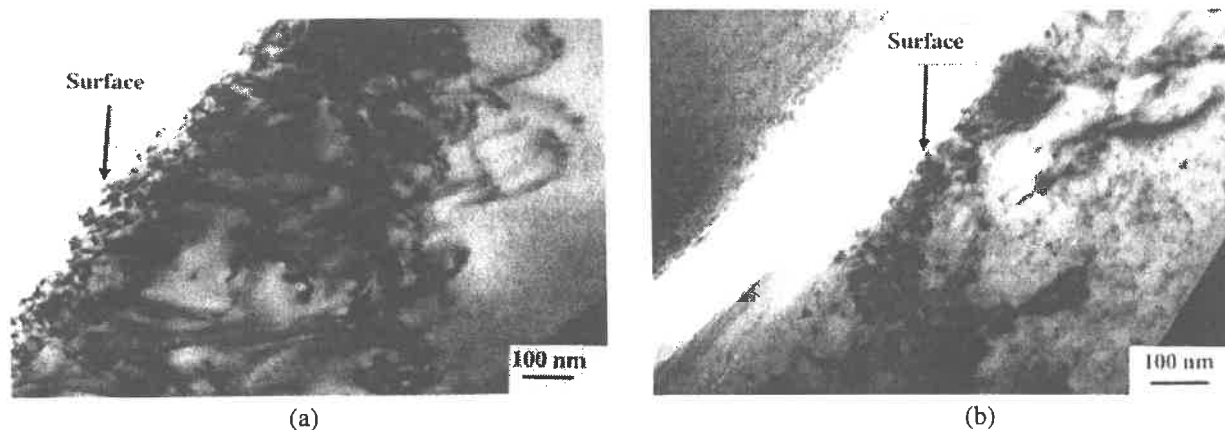


Fig.5 Subsurface structure in alumina after ductile-mode grinding. (Table speed 1 m/min): (a) The 1 μm -grained alumina; (b) The 25 μm -grained alumina.

distinguished easily there. The depth of the second zone was also dependent on table speed and varied from 150 nm to 500 nm.

3.3 Microcracks

Microcracks were observed only in the vicinity of pit edges, which neither radial nor lateral cracks. The microcracking-free behaviour in the first layer can be explained by the nucleation of sufficient number of twin and slip systems that satisfy von Mises criterion [4]. The second zone had less than five independent slip and twin systems and thus microcracking should be highly possible there. However, this is inconsistent with the above experimental observations. Such phenomenon can be explained by the possibility of microcracking due to the pile-up of dislocations on an active slip plane. The threshold effective resolved shear stress (τ_s) for cracking could be determined by [5]:

$$\tau_s^2 = \frac{3\pi}{8} \left[\frac{\gamma\mu}{(1-\nu)L} \right]$$

where γ is fracture surface energy, μ is shear modulus, ν is Poisson's ratio, and L is the length of pile-up. Considering the length of pile-up of 150-500 nm (i.e. penetration depth of dislocation in grinding), we can see that for such a small pile-up length the nucleating of microcracks usually does not occur. It is also worth to note that the interaction of different slip systems was a rare event here and could not create microcracks.

4. Conclusions

- (1) Ductile-mode of material removal was obtained in ultra-precision grinding of different types of alumina, featuring mirror surfaces with Rms roughness in the range of 30 nm to 90 nm.
- (2) The nature of pits on ground surfaces is related to pores in the bulk material.
- (3) The subsurface structure of alumina after ductile-mode grinding is composed of a layer with high dislocation density in the surface vicinity followed by a layer with much lower dislocation density.
- (4) The ductile-mode of material removal is possible due to initiation of more than five independent slip and twin systems in the first layer. The possible explanation of the absence of microcracks in the second layer is the small length of pile-ups that can not create sufficient stress for microcracking.
- (5) No radial and lateral microcracks were detected in the whole subsurface area after grinding.

References

1. Zarudi and L. C. Zhang, in: *Advances in Abrasive Technology*, L. C. Zhang and N. Yasunaga, ed. 33 (World Scientific, Singapore, 1997).
2. H. Suzuki, N. Wajima, M. S. S. Zahmaty, T. Kuriyagawa, K. Syoji, in: *Advances in Abrasive Technology*, L. C. Zhang and N. Yasunaga, ed. 116 (World Scientific, Singapore, 1997).
3. I. Zarudi and L. C. Zhang, Proc. of 10-th ASPE Annual Meeting, 165 (1995).
4. R. Von Mises, *Math. Mech.*, **8**, 161 (1928).
5. T. Hagan, *J. Mat. Sci.*, **14**, 2975 (1979).

IN-PROCESS EDM TRUING TO GENERATE COMPLEX CONTOURS ON METAL-BOND, SUPERABRASIVE GRINDING WHEELS FOR PRECISION GRINDING STRUCTURAL CERAMICS*

Mark A. Piscotty, Pete J. Davis, Theodore T. Saito, Kenneth L. Blaedel, Lee Griffith
Lawrence Livermore National Laboratory
7000 East Ave Livermore, CA 94551 USA
510-422-2052 / 510-423-1460 fax
markp@llnl.gov

1. INTRODUCTION

The demand and use of precision grinding of structural ceramics continue to increase as the worldwide advanced ceramic industry surpasses \$20 billion in sales [1]. Included in this industry are engineering structural ceramics, electronic ceramics, bioceramics and others. These materials are used in applications such as engine components, casting and extrusion dies, bearings, medical implants, nozzles, thermal insulators, and more. Along with the variety of ceramic applications comes a broad range of precision requirements, which in turn leads to various required processes to accommodate a spectrum of specifications. A process for grinding ceramic components to micrometer tolerances was employed and further developed at Lawrence Livermore National Laboratory for two separate grinding projects.

2. PROCESS DESCRIPTION

The grinding methodology was developed for a variety of ceramic grinding processes including creep-feed and cylindrical grinding. Figure 1 shows a schematic with the major components of the grinding machine used to creep-feed grind flat ceramic components. An in-situ EDM rotating graphite electrode is used to profile metal bond grinding wheels and both the electrode and the grinding wheels are mounted on air-bearing spindles [2]. The rotating electrode has separate truing surfaces, for coarse and fine profiling. In turn, the profile that is imparted onto the grinding wheel may have multiple, complex features for coarse stock removal and finish detailing.

2.1 Creep-Feed Grinding

One application which greatly benefits from adoption of this methodology is creep-feed grinding of precision grooves in flat BeO substrates. Figure 2

*This work was performed under the auspices of the US Dept. of Energy by LLNL under Contract W-7405-Eng-48

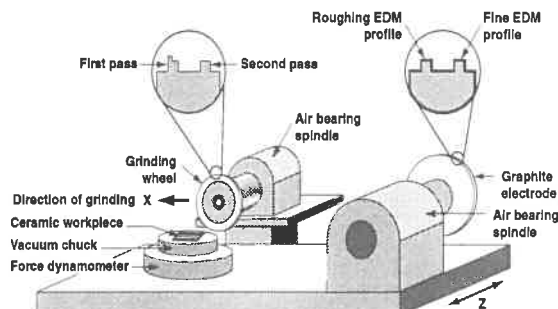


Fig. 1. Creep-feed grinding process schematic with in-situ EDM wheel profiling.

shows a photograph of a ground 1 cm x 4 cm x 0.2 cm BeO component. Most tolerances are $\pm 1 \mu\text{m}$, including the periodic spacing among the grooves. In addition to meeting dimensional specifications, it is necessary to maintain surface roughness and sub-surface damage below levels that significantly affect the material properties and component performance [3].

Figure 3(a) shows a close-up photograph of the profiled grinding wheel approaching four BeO workpieces held in a vacuum chuck, which is mounted on a 3-axis, strain-gage force dynamometer. The machine tool is temperature controlled to $\pm 0.5 \text{ C}$ and air-bearing spindles support and drive the grinding wheel and graphite electrode.

Low pressure aqueous grinding fluid is used with the option of applying high pressure fluid to

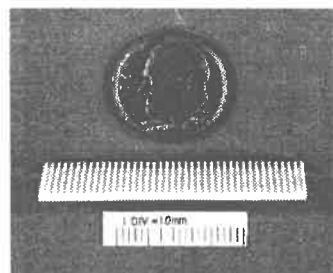


Fig. 2. 4 cm x 1 cm x 0.2 cm BeO substrate.

continuously clean the wheel surface. Figure 3(b) shows the graphite electrode about to profile the metal bond, grinding wheel. For this particular process, a CBN grinding wheel, mounted on the same spindle as the diamond grinding wheel, is used to profile the graphite electrode. Table 1 provides typical process parameters for the creep-feed grinding.

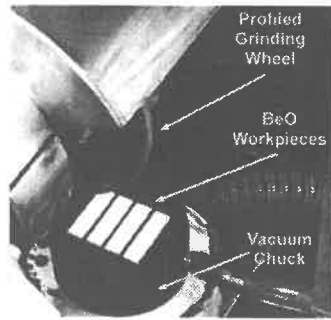


Fig. 3(a). Creep-feed grinding of BeO.

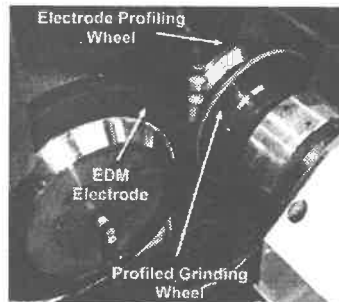


Fig. 3(b). EDM profiling of grinding wheel.

Table 1. Creep-feed process parameters.

Parameter	Description
Grinding wheel	SD1000, N100 M, 18 cm Ø
Workpiece	BeO
EDM electrode	Poco Graphite (1 µm grain size)
Wheel speed	105 m/sec
Feed rate	7.5 cm/min

2.2 Cylindrical Grinding

Another application that utilizes this process methodology is cylindrical grinding of structural ceramic engine components. These multi-featured components have typical tolerances down to a micrometer. Figure 4 shows a photograph of one of the zirconia components.

Some of the critical features to note regarding

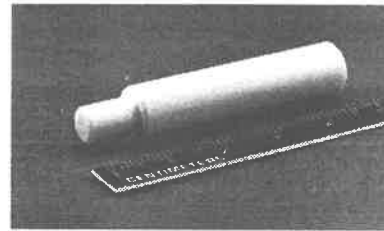


Fig. 4. Ground zirconia component.

this geometry are cylindricity to within 1.5 µm, a specified radius of curvature on the right side of the component, a chamfer edge on the left side and a minimum shoulder internal corner radius.

The process procedure is similar to that for creep-feed grinding except that the workpiece is rotated in a spindle mounted collet, instead of mounted on a traversing table. In this case, in-situ profiling of the EDM electrode is performed using a single-point tool mounted on the machine. The profiled electrode is used to EDM profile the grinding wheel.

Figure 5 shows a photograph of the grinding machine and shown are the grinding wheel spindle and the workpiece spindle that rotates the workpiece and the EDM electrode. Also shown are the ceramic workpiece and the single point tool used to profile the electrode. Table 2 presents the nominal process parameters used in this particular application.

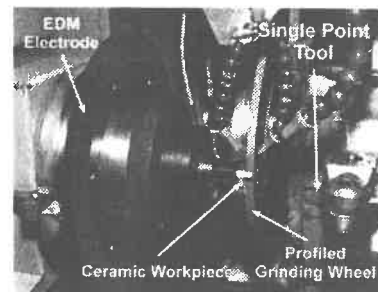


Fig. 5. Cylindrical grinding machine tool.

Table 2. Nominal process parameters for cylindrical grinding using EDM profiled grinding wheels.

Process Parameter	Description
Grinding wheel	15-25 µm, cast iron bond, 30 cm Ø
Workpiece	Al ₂ O ₃ , Si ₃ N ₄ and Al ₂ ZrO ₂
EDM electrode	Poco Graphite (1 µm grain size)
Wheel speed	25 m/sec
Workpiece speed	150 rpm

3. PROCESS ANALYSIS

3.1 Grinding Wheel Wear

Understanding and controlling wheel wear is an essential factor in determining both the level of achievable precision and the economics of the grinding process [4]. To facilitate this understanding, an instantaneous grinding ratio (volumetric material removed/volumetric wheel wear) test procedure was developed to examine the changes in grinding ratio over time. To circumvent the problem of directly measuring small changes in wheel dimensions, measurements of witness grinds were used.

Figure 6 shows a schematic of the instantaneous grinding ratio method using cylindrical grinding. A workpiece is rotated in a spindle and a grinding wheel grinds with only half of the available bond surface, leaving the unused portion intact. Shallow witness grinds, with both the 'unused' and the worn portions of the wheel, are performed on available workpiece areas after each successive grind. This produces a measurable artifact for step height differences between the worn and unused wheel surfaces.

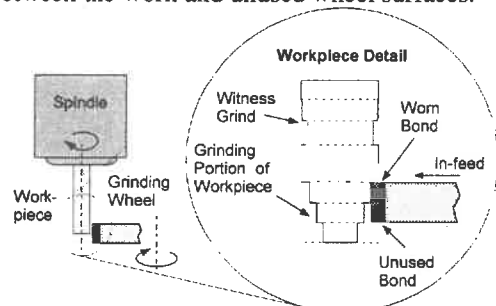


Fig. 6. Instantaneous grinding ratio schematic.

Figure 7 shows an image of a typical ground workpiece from such a test. On the left are shown the grinding steps and on the right is one of the witness grinds used to measure wheel wear. Figure 8 is a plot of data obtained for a particular grinding wheel for different workpiece materials.

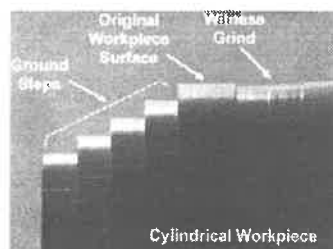


Fig. 7. A typical workpiece ground during an instantaneous grinding ratio test.

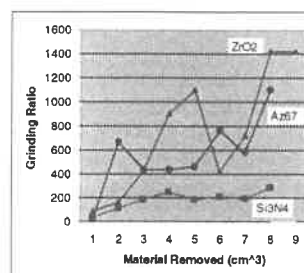


Fig. 8. Instantaneous grinding ratio data for various materials. Wheel: 15-25 μm , cast iron bond.

grinding conditions shown in Table 1 and a wheel traverse speed of 5.1 cm/min. This workpiece showed a significant number of cracks between and through grains to depths of 10 μm or more. A sample ground at 0.7 cm/min showed essentially no cracks, but had a larger number of dislocations near the surface. We surmise that the more catastrophic damage of cracks may significantly affect the thermal and mechanical properties of the component.

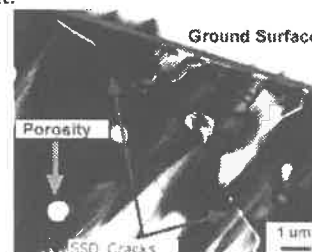


Fig. 9. TEM image of a ground BeO surface with wheel traverse speed of 5.1 cm/min.

3.2 Sub-surface Damage

Investigation of the presence of residual damage, including sub-surface damage (SSD) [5], and its effect on workpiece material properties is being investigated. Transmission electron microscopy (TEM) analysis was used to examine the SSD. Figure 9 shows a typical TEM of a BeO surface using

Another type of grinding test was conducted to investigate the effect of residual damage on the modulus of rupture (MOR) values of zirconia-alumina ($\text{ZrO}_2\text{-Al}_2\text{O}_3$) bars. Three groups of bars were machined, then tested in a 4-point flexure [6]. The first group was loose abrasive polished (0.5 μm diamond), the second was ground with a 25 μm diamond abrasive, bronze bond wheel, and the third was ground with a 100 μm diamond abrasive, bronze bond wheel. Figure 10 shows a plot of the

MOR values for the three cases and indicates that the polished and the fine ground parts had statistically similar MOR values, while the coarse ground parts exhibited a reduction of MOR by a factor of approximately 3.

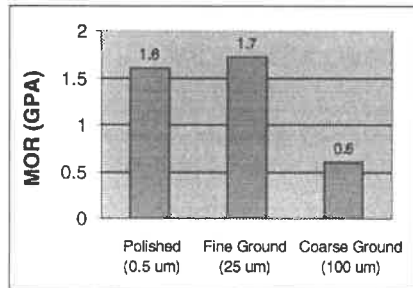


Fig. 10. MOR values as a function of grit size. $\text{ZrO}_2\text{-Al}_2\text{O}_3$ MOR bars.

4. DISCUSSION

The economic advantages of instituting an on-machine EDM profiling process for precision grinding can be substantial for a number of reasons. Integrating the entire process on one machine reduces operator set-up and tear-down times over multiple-machine procedures. Using a single machine tool reduces capital equipment expense and decreases required floor space. However, one of the largest economic incentives of such a procedure is that performing all the machining and grinding on one machine can significantly improve part accuracy leading to improved yield. Equally important is the versatility to generate complex profiles in grinding wheels, which can reduce the number of passes required to grind a given component.

Control of wheel wear is essential to controlling the accuracy of the grinding process. Traditional grinding ratio tests typically require substantial grinding, with a single 'integrated' grinding ratio value as the result. The instantaneous grinding ratio tests discussed here allow the user to capture wheel wear information throughout the life of the wheel. The data shown in Figure 8 all show a sharp increase in grinding ratio after the first few grinding steps. The initial low grinding ratios are the result of the softer, recast layer generated on the wheel surface during EDM profiling, which wear quickly relative to the bulk bond matrix. To maintain dimensional accuracy, the user must be aware of how the wheel wears over time and optimize the wheel profile and EDM parameters to account for the recast layer.

The residual SSD imparted in the ceramic workpiece is often as important to the customer as is compliance with geometric specifications. As higher performance demands are placed on ceramic components, changes of 10 – 15% in mechanical properties can become crucial. In one test, we noted a 3 times decrease in modulus of rupture values for coarse ground parts over polished or fine ground parts. This has significant implications regarding component reliability and time between failures. Further research will continue at LLNL to investigate the thermal and mechanical SSD effects on ground ceramic components.

5. SUMMARY

A generalized precision grinding process is under development at Lawrence Livermore National Laboratory to use EDM profiled, metal-bond, superabrasive grinding wheels to economically fabricate ceramic components. The two applications reported here include creep-feed grinding of 0.5 mm wide grooves in BeO substrates and cylindrical grinding of various ceramic components with micrometer tolerances. Resultant sub-surface damage and grinding ratios were investigated to optimize grinding conditions for maximum yield and wheel life.

6. REFERENCES

- ¹ Mayer, J.E., "Structural ceramics tell you their optimum grinding condition," *Abrasives Magazine*, pp 14-18, June/July 1997.
- ² Suzuki, K., "On-machine electro-discharge trueing for metal bond diamond grinding wheels for ceramics," NIST Special Publication 847, Editor – Said Jahanmir, July 1993.
- ³ Green, D.J., "Critical microstructure for microcracking in $\text{Al}_2\text{O}_3\text{-ZrO}_2$ composites," *J. Am. Ceram. Soc.*, 65, pp 610-614, 1992.
- ⁴ Mindek, R.B., "A study of the wheel wear and thermal limitations in creep-feed grinding...", PhD Dissertation, Univ. of Connecticut, 1996.
- ⁵ B. Zhang, "Subsurface evaluation of ground ceramics," *Annals of CIRP*, Vol 44/1/1995.
- ⁶ "Standard test method for flexural strength of advanced ceramics at ambient temperature," ASTM Standard C 1161-90, December, 1990.

Figure correction for rotary asymmetric error by linearly arranged segmented pressurizing polishing machine

Ken-ichiro Horio, Toshio Kasai, Atsuhiko Matsuse

Saitama university

255 Shimo-okubo, Urawa, Saitama, 338, Japan

1. Introduction

For example, in a field of optical components or semiconductor wafers there is no bounds for demand of high performance. In some cases we will have to mass-produce those high quality articles which has both smooth surface and high form accuracy. In the mass production yield rate of these should be rather high. One solution for dealing with those is advancing the final figuring process. However it is not effective nor economical in some cases. The other solution is adding a special figure correction process to the ordinary final process. Of course, in this case, you should have some degree of yield rate from the ordinary final process and have some inferior articles just below the border. Unless this condition is satisfied, the adding process is not effective nor economical. Our aim of the work is for this kind of figure correction technology.

2. Advantage of segmented tool method

In the previous work[1] we proposed the effectiveness of using multiple small tools in the polishing for figure correction. Fig.1 shows the comparison between (usual) small tool method and segmented tool one. In the small tool method you should scan a tool over the sample and control dwelling time at each position to compensate material removal amount. On the contrary, in the segmented tool method the compensation is done by supply of pressure to each piston. There are two

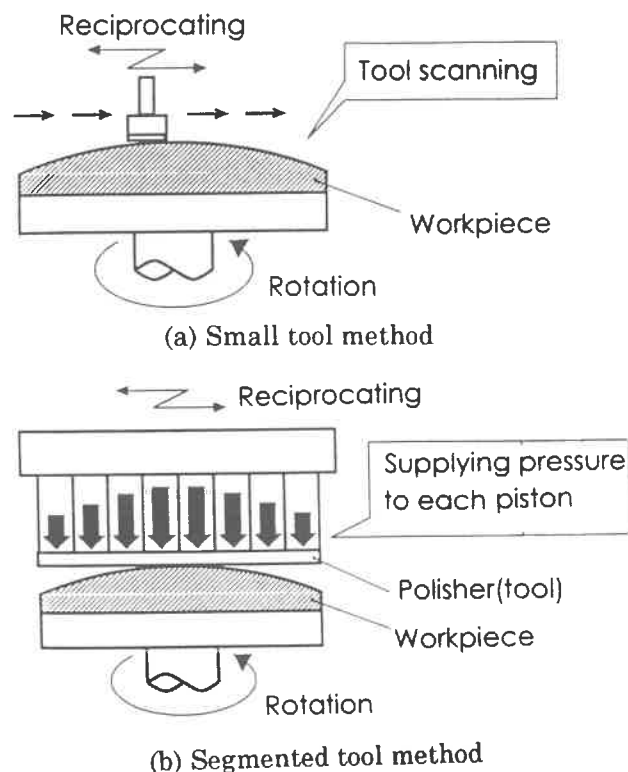


Fig.1 Comparison of principle between small tool method and segmented tool one

way of the compensation; one is by amount of pressure and the other is by time with constant pressure. though the time is easier to be controlled.

Because you can expect several times efficiency with 10 small tools in compare of single tool. We also stated linearly arranged segmented tools is more practical or effective than two-dimensionally arranged segmented tools which has some difficulty in pressuring or supplying abrasives.

On the contrary linearly arranged tool is hard to compensate the rotary asymmetric error because the apparatus should be moved with more than two degrees of freedom (if you use two-dimensional, one degree is enough) and the one degree tends to be rotary movement in easier design. So in the previous work results were limited to rotary symmetric figures.

3. Strategy for correction of asymmetric error

In this work, we tried to treat asymmetric error by rotary movement. Fig.2 shows the principle. In the figure correction process the radial position is always detected by the encoder and orders of pressuring were sent dependent on the position. The figure correction is based on dwelling time control of each tool. The strategy is integration of material removal by each tool and consideration of overlapping with each other. In this case the difference of radial position of each tool also should be considered.

4. Asymmetric Figure Correction by Prototype Machine

We made a prototype apparatus to demonstrate and prove the idea. Fig.3 shows the configuration of the system. It has ten segmented tools which is linearly arranged and can be moved linearly as one body along the same direction of the line. Workpiece is fixed on the rotary table with a rotary encoder. Fig.4 shows the tool arrangement and its capability to cover an area over the sample (experimental result). We developed a program to control each electro-magnetic valve for pressurizing a piston by the signal from the encoder and the information about the figure deviation

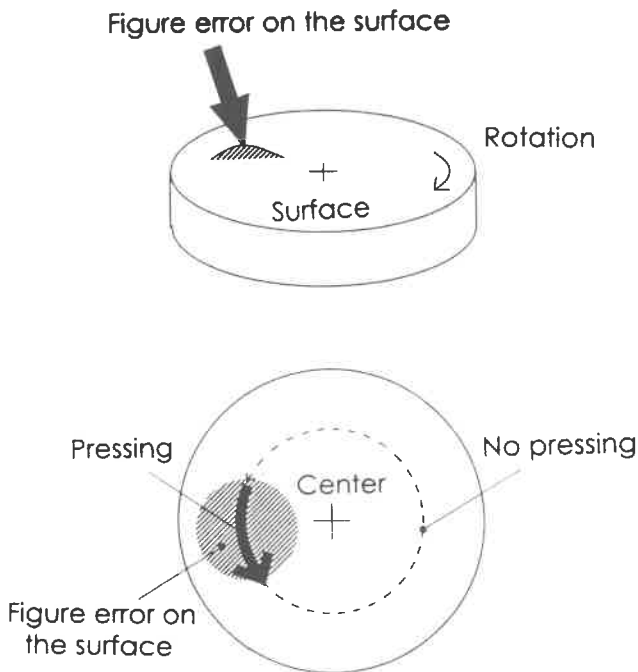


Fig.2 Principle of the figure correction for asymmetric error

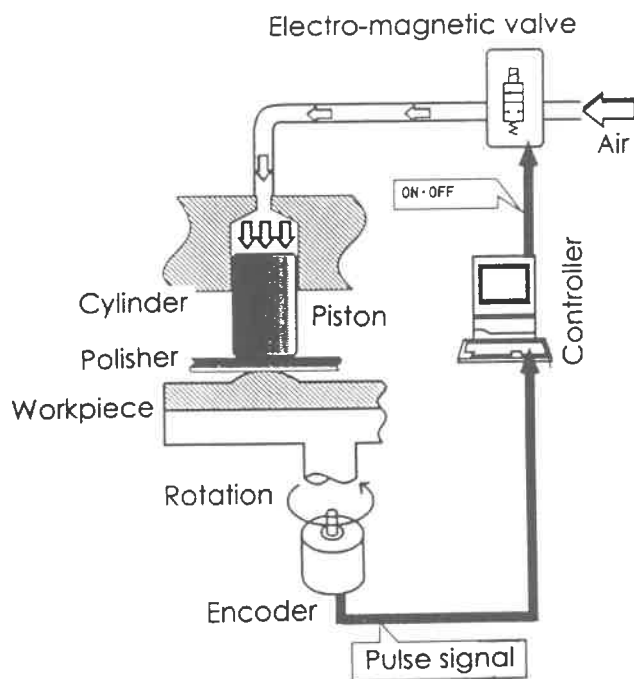


Fig.3 Conceptual figure of the proto-type apparatus

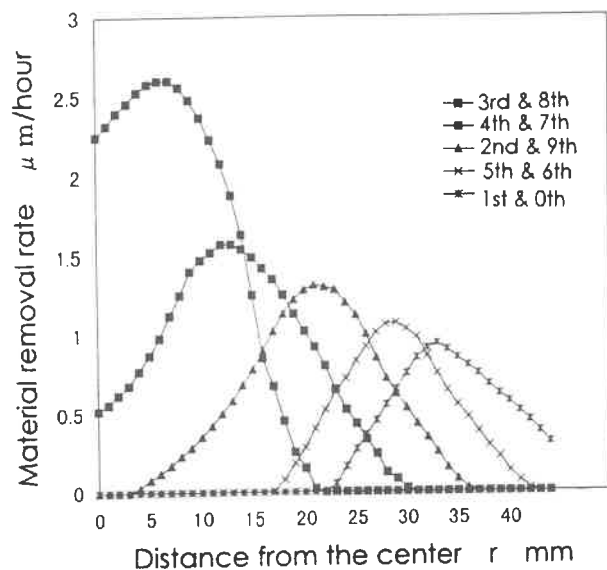
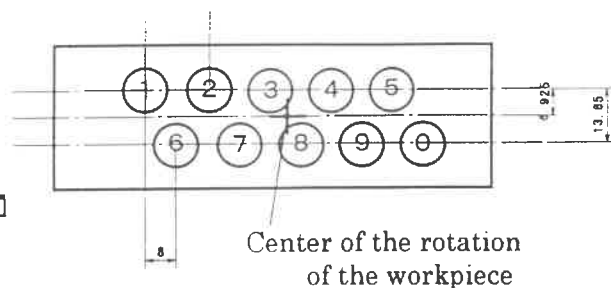


Fig.4 Arrangement of the tools in the prototype machine and its covering area

(where should be removed).

For the example, we tried to get a cylindrical concave which is rotary asymmetric with this rotary movement apparatus. Fig. 5 shows the result. Work material was BK7 (optical glass). Though the initial shape of the sample was spherical convex, we could get a shape like cylindrical concave by polishing combined with cerium dioxide powders and artificial leather polisher. Fig.6 shows the sectional shape along the two main directions (generator and diameter) of the cylinder. You can find difference of those.

5. Conclusion

Though our prototype is mechanical polishing apparatus, the idea can be applied to other complicated process as it is under rule of dwelling time only. The applicable area we think are manufacturing of aspheric optical (glass) components, planarization during semiconductor process, or precision molding die for plastic optical components and flatness improvement in wafers for

semiconductor or quartz crystal.

- [1] Horio K., Kasai T., 1993, Figure Correction by Segmented Pressurizing Polishing Machine, Proc. 1993 ASPE Annual Meeting: 268-271

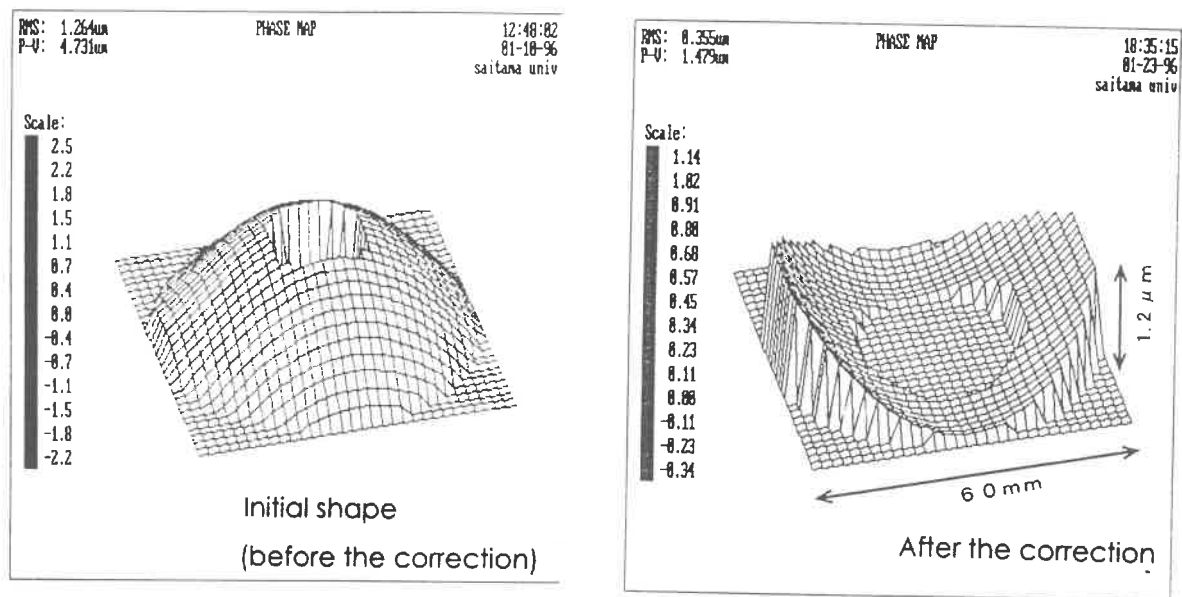


Fig.5 Example of the figure correction to a cylindrical concave

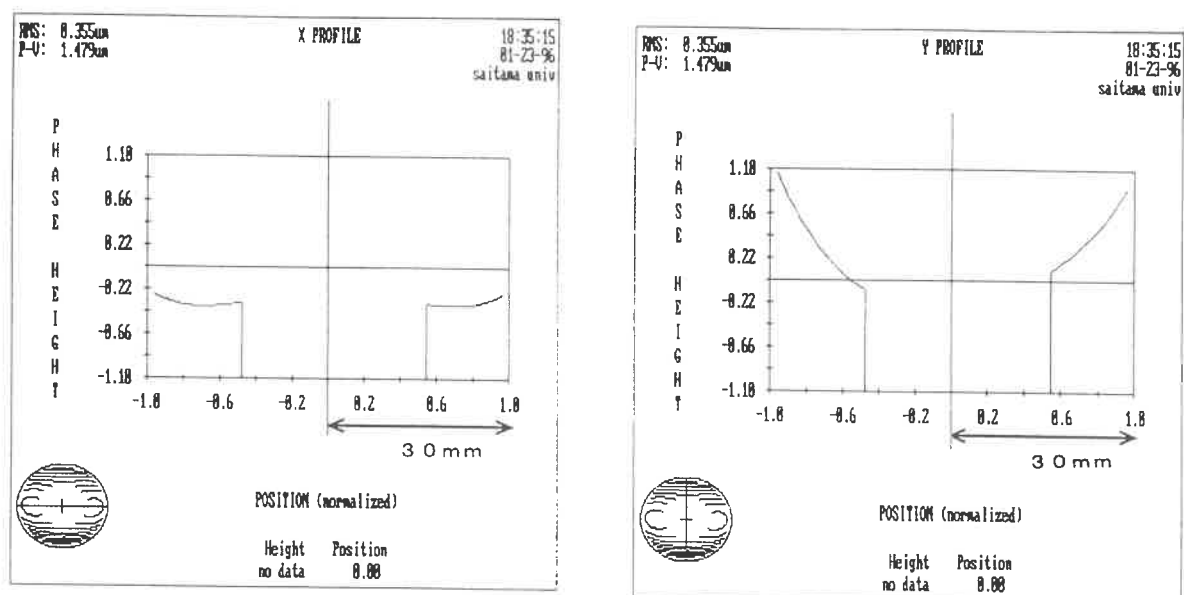


Fig.6 Sectional shapes of the corrected sample

Fabrication of new collimating optical fibers using glass slit-microferrules

Kunio KOYABU and Fumikazu OHIRA

NTT Opto-electronics Laboratories, Nippon Telegraph and Telephone Corporation
3-9-11, Midori-cho, Musashino-shi, Tokyo 180 Japan

1. Background

Efficient input of light beams into optical fibers and output from fibers is important for optical communications. Methods where optical lenses are placed in front of the end surfaces of optical fibers have been widely used to improve input and output efficiency. However, it is difficult to align lenses to optical fibers and the parts making the joint are too big. To overcome these problems, a new structure in which a piece of graded-index (GI) fiber is fused and connected to the end of a single-mode (SM) fiber, has been proposed and evaluated[1]. This structure overcomes the problems of the difficult fabrication and the excessive volume, but creates new problems in that the optical characteristics are affected by the heating process and the irregular length of the GI fibers from the cleaving. Therefore, we propose a new collimating optical fiber that uses glass slit-microferrules. We have fabricated such a fiber and evaluate its optical characteristics.

2. Structure and principle of the collimating optical fiber

In the proposed collimating optical fiber (C-fiber), a piece of GI fiber with a regular length is placed in contact with an end surface of a SM fiber inside a glass slit-microferrule (Fig. 1), and they are fixed together with ultraviolet (UV) glue. A C-fiber transforms a divergent light beam into a collimated beam that exits the end surface of the C-fiber. To collimate light beams, the length of a piece of GI fiber should be $(2n+1)p/4$, where p is the period of a sine-curve and n is an integer, because a light beam from a SM fiber propagates along the trajectory of a sine-curve in a GI fiber.

3. Fabrication process and advantages of C-fibers

The fabrication process of the C-fiber is shown in Fig. 2. First, the coat of the SM fiber is removed and the bare part of the SM fiber is cleaved to a length of 6-7 mm. Then, the bare part of the SM fiber is set into a glass slit-microferrule. The slit-microferrule is a glass tube with a slit formed by dicing, and it has an outer-diameter of 0.25 mm, an inner-diameter of 0.126 mm and a length of 10 mm. Next, a piece of GI fiber is set into the opposite side of the glass slit-microferrule from the SM fiber, and UV glue is applied to the interface between the two fibers through the slit. The glue is then hardened by exposure to UV light, fixing the two fibers and the glass slit-microferrule together. After that, the GI fiber and the glass microferrule are polished until the GI fiber reaches the desired length.

This fabrication process offers the advantages of a short process with only four steps, a simple fabrication without two-fiber positioning, no damage to the optical characteristics from a heat treatment and good reliability in the optical characteristics because the GI fibers can be polished.

4. Experimental results

In the design of C-fibers, it is important that the periods of the sine-curves are clear when light beam are propagating inside the cores of GI fibers. Figure 3 shows the relation between the length of GI fibers with cores of 50 and 80 μm and the divergent angle of the outgoing light beam. The period of the sine-curve for 50 μm is 1.1 mm and for 80 μm it is 1.4 mm. Table 1 shows the specifications of fabricated C-fibers with 50- μm and 80- μm cores, and Fig. 4 shows a fabricated C-fiber with a 50- μm core.

The relation between the traveling distance and the light-beam diameter when a beam leaves a C-fiber is shown in Fig. 5. The outgoing beam diameter ($1/e^2$) of a C-fiber with a 50- μm core is less than 0.3 mm at a traveling distance of 6 mm and the divergent angle is 2.5 degrees. Our experimental results showed that the divergent angles of these C-fibers were about one-quarter those of conventional SM fibers. Results for the C- fibers with 80- μm cores were better than with the 50- μm cores, which suggests that the larger the core of a GI fiber, the better the optical characteristics.

Figure 6 shows the contours and the distributions of an outgoing light beam from a C-fiber with a 80- μm core at traveling distance of 0.5 mm and 5.5 mm. At 0.5 mm, the contour consisted of various circles with the same center position and a good Gaussian distribution was obtained. The contours and the distribution at 5.5 mm were almost as good. Thus, fabricated C-fibers have good optical characteristics that change little while traveling a certain distance.

Figure 7 shows the relation between the C-fiber interval and optical coupling loss when two C-fibers were placed with various positioning errors. For small C-fiber positioning errors, the optical coupling loss simply worsened as the intervals of the C-fibers became larger. For some of the larger C-fiber positioning errors, though, coupling losses had local peaks while one of the C- fibers moved. It appears that for small fiber positioning errors, the input (detection) C-fiber will always be placed in the outgoing light beam at any intervals, however, for larger positioning errors, the input C-fiber may not be placed in the outgoing light beam for short intervals, but comes into the outgoing light beam as the interval widens.

With the experimental C-fiber, we achieved non contact connecting with 3-dB coupling losses for 0.5-mm intervals between C-fibers with a 50- μm core and for 1-mm intervals with an 80- μm core.

5. Conclusions

We have proposed and fabricated C-fibers made using a glass slit-microferrule. In this structure, a piece of GI fiber with a regular length is connected to the end of an SM fiber inside of a glass slit-microferrule and they are fixed together with UV glue. The fabricated C-fibers had good optical characteristics because of the short and easy fabrication without heating treatment. Non contact optical connection using the C-fibers was achieved.

Reference

- [1] W. L. Emkey and C. T. Jack, " Analysis and Evaluation of Graded-Index Fiber-Lenses ", Journal of Lightwave Technology, VOL. LT-5, NO. 9, Sep., pp. 1156-1164, 1987.

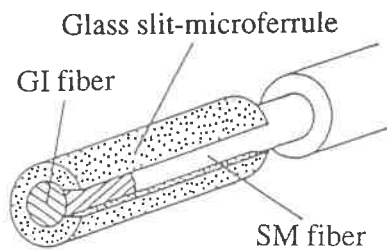


Fig.1 Structure of collimating optical fiber using glass slit-microferrule

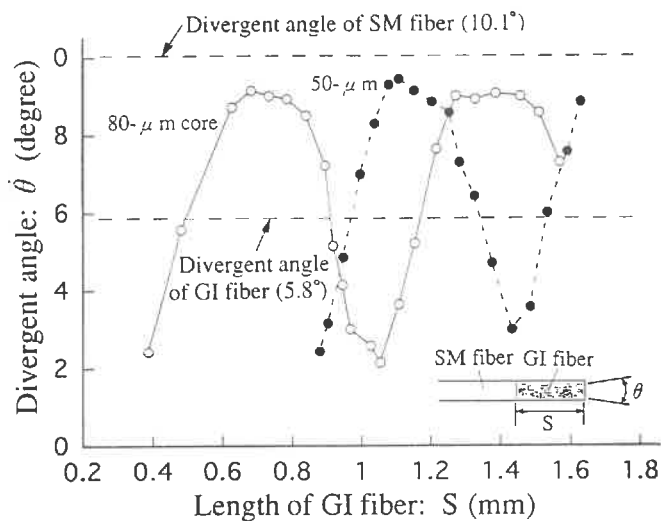


Fig. 3 Relation between length of GI fiber and divergent angle of outgoing light beam

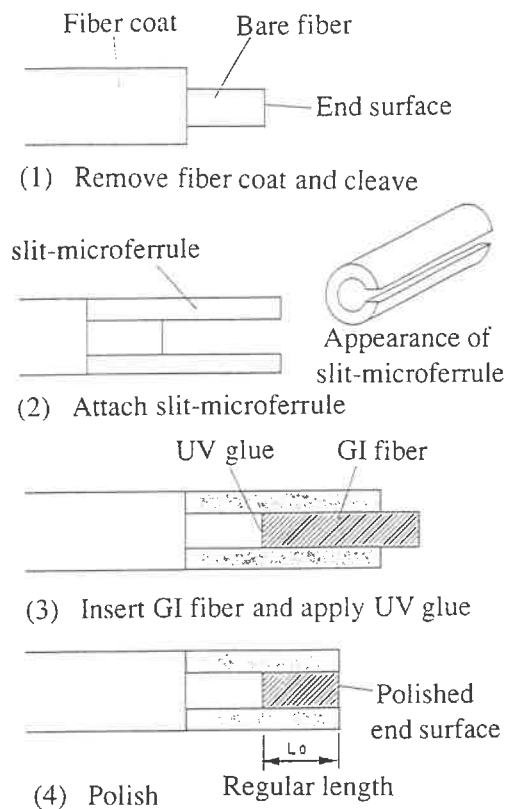


Fig. 2 Fabrication process of C-fiber

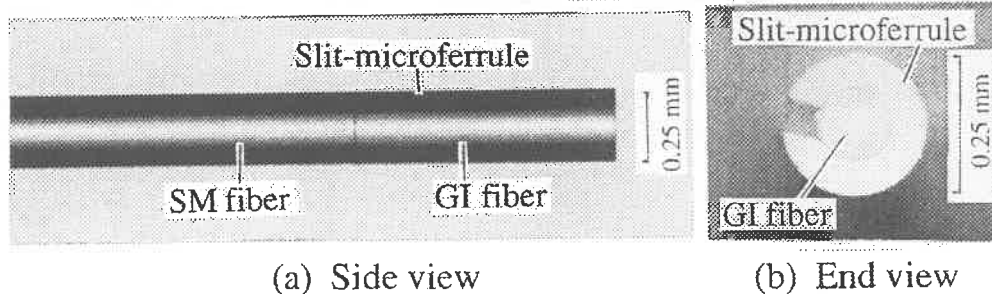


Fig. 4 Fabricated C-fiber

Table 1 Specifications of fabricated C-fibefs

Core diameter of GI fiber	Length of GI fiber	Core diameter of SM fiber	Outer-diameter of C-fiber
50 μ m	0.85 mm	6 μ m	0.25 mm
80 μ m	1.1 mm		

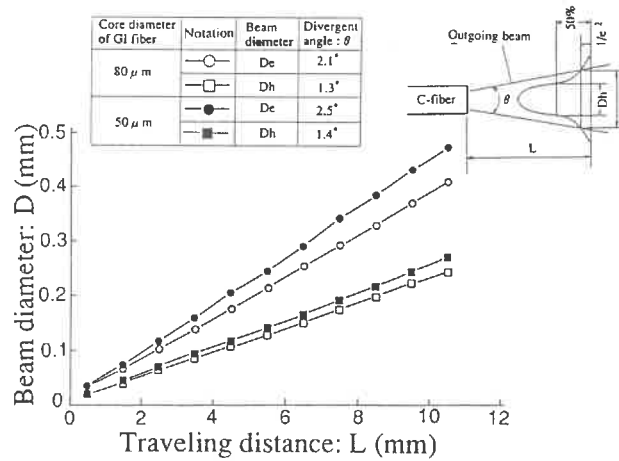


Fig. 5 Relation between beam diameter and traveling distance of outgoing beam from C-fibers

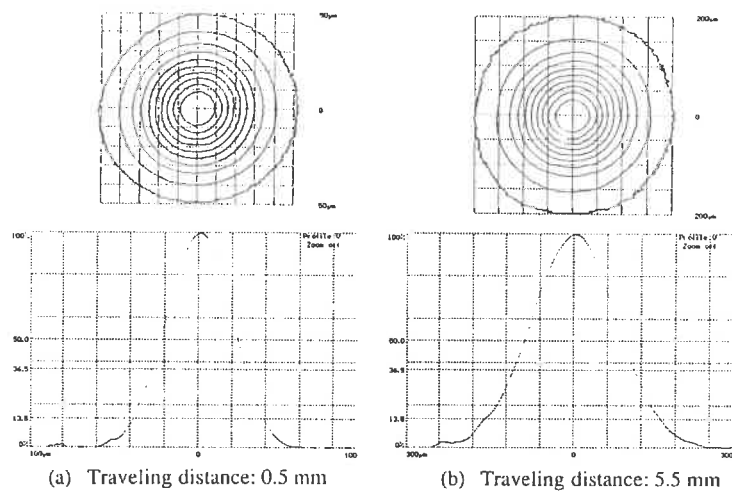


Fig. 6 Contour and distribution of a light beam from a C-fiber with an 80- μm core

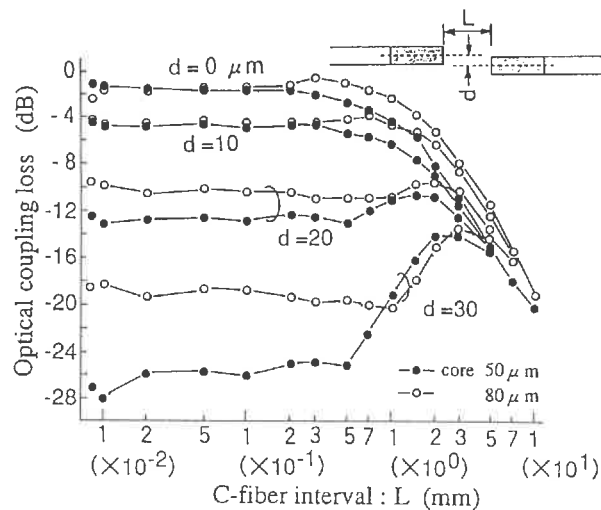


Fig. 7 Relation between C-fiber interval and optical coupling loss

ROTATING LAPPING MACHINE WITH A PNEUMATIC BEARING SUPPORTED LAP

Gabriel Ascanio⁽¹⁾, Carlos J. Cava⁽¹⁾, Roberto U. Reséndiz⁽²⁾, Ricardo Chicurel⁽³⁾ and Rigoberto Nava⁽¹⁾

(1) Centro de Instrumentos, UNAM, A.P. 70-186, Mexico, DF 04510

(2) DIBYASA, Indiana 260-704, Cd. de los Deportes, Mexico, DF 03710

(3) Instituto de Ingeniería, UNAM, Mexico, DF 04510

Abstract

The development of a prototype rotating lapping machine is presented. In order to improve the flatness and finish of the workpiece, a pneumatic bearing has been used to support the lap. Many factors such as stability and high load capacity were taken into account to design the bearing. With respect to the kinematics, lapping is characterized by the type of relative motion of the lap and work surface. The lap rotates around the machine axis. The workpiece holder can also rotate resulting in a rotational and translational relative motion.

Introduction

Lapping is regarded as the oldest method of obtaining a fine finish, and is extensively used in the engineering world in applications such as the finishing of gauge blocks and flats for checking production parts. Lapping is an abrasive process carried out by means of loose abrasive in which the role of the cutting points is played by grains finding momentary support by the laps. Lapping has four main purposes: to obtain a superior accuracy as to dimensions, to correct shape imperfections, to obtain smooth polished surfaces and to improve the fit between surfaces. The purpose of using a pneumatic bearing is to damp out vibrations that can have a negative effect on the surface quality. No reports of the applications in lapping machines were found in the literature so that the concept appears to be novel.

Description of the system

In the rotating lapping machine discussed in this paper, a rigid structure supports the mechanism and the drive system, which consist of the lap, the pneumatic bearing and the motor. The lap used in this equipment is a 12 inches diameter and 1 inch thick plate made of ASTM 30 soft cast iron [1] and the conditioning rings are 4¼ inches cylinders made of AISI 4140 steel. The composition of cast iron is 3.25% C and 2% Si with Brinell Rockwell Hardness of 80. A soft material is used in accordance with the fact that the lap should always be softer than the material to be lapped in order to avoid charging the workpiece, and cutting of the lap itself [2]. Harder laps will often cause glazing and scratching while softer laps cause a loss of flatness and parallelism and will produce grayer finishes. Care was taken to make the plate heavy enough so that it would not distort in use. Workpieces to be lapped are supported by three holders mounted over

the lap at 120° between holders. The frictional force can be varied by using weights mounted on the workpiece holders. In this prototype, the abrasive and liquid are manually supplied.

The pneumatic bearing consists of two complementary conical parts. This geometry allows axial and radial loads in the same plane. The pneumatic bearing has two elements: one static (lower) and the other floating (upper) [3]. The clearance between the two elements is 2.8×10^{-5} m, when pressurized air is introduced at 490 kPa.

The drive system consists also of a ½ hp 4 pole induction motor coupled to a gear speed reducer with 5.1 ratio, so that the maximum rotational speed is 350 rpm. The lapping speed is controlled by means of a frequency changer which is directly connected to the motor. Normally, lapping equipment operates at a fix lapping speed, which is of the order of 250 rpm [2], however, in this prototype testings at different speeds will be possible. Othe operating conditions that will be varied are the abrasive size and lapping pressure to determine their effect on the stock removal rate [4,5]. Figure 1 shows a diagram of the rotating lapping machine and the conical air bearing.

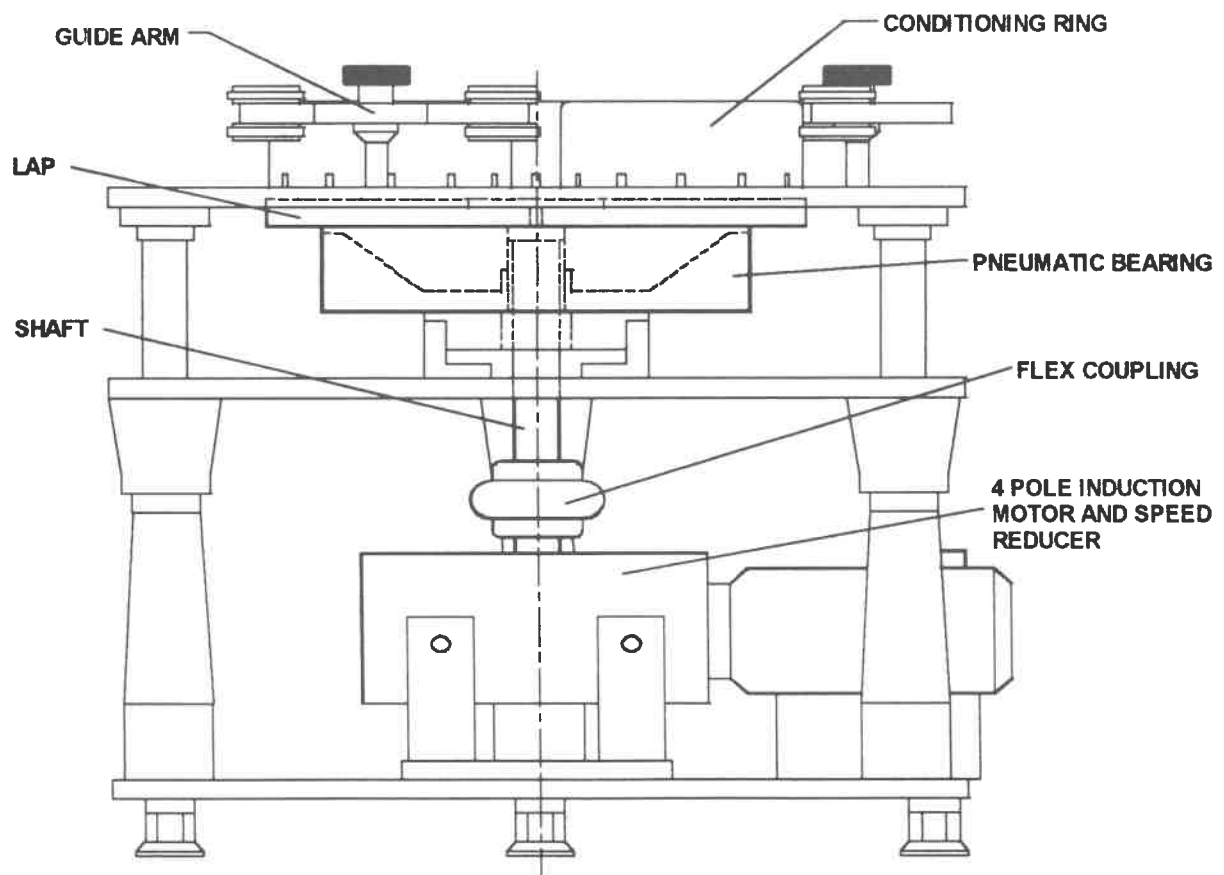


Figure 1 Rotating lapping machine with a pneumatic bearing

Kinematics of lapping machine

In the present lapping machine, the workholders rotation is induced by the drag exerted by the lap. A fundamental problem arises is the calculation of the rate of such induced rotation as a function of the lap's angular velocity.

Following Matsunaga [5], the problem can be solved by assuming a constant pressure distribution, P , and friction coefficient, μ , over the contact area between the workpiece and the lap. Then the friction force per unit area at any point is μP and its direction is along the relative velocity vector.

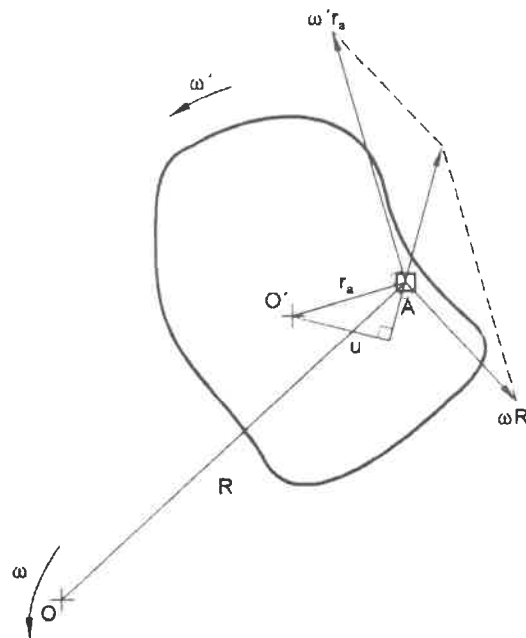


Figure 2 Velocity diagram of an arbitrary shape

Figure 2 is a diagram in which O represents the lap's axis and O' the workholder's axis. A workpiece of arbitrary shape is shown. The angular velocity of lap and workpiece are ω and ω' respectively. Other symbols are defined in the figure. Any point of the workpiece such as A has a velocity relative to a coincident point on the lap which is made up of two components: ωR in a direction perpendicular to OA and $\omega' r_a$ in a direction perpendicular to $O'A$. The resultant is in the direction of the friction force at that point. u represents the moment arm of such friction force about O' . The total moment on the workpiece is obtained as

$$M = \iint_A \mu P u dA$$

In the above equation, u may be expressed in term of appropriate coordinates of the point A and the velocities ω and ω' . Upon carrying out the integration and setting $M = 0$, the desired relation between ω and ω' is obtained.

Test setup

In order to determine the effect of lapping pressure on the stock removal rate, five ½ inch diameter test pieces made of AISI 1045 steel quenched and tempered will be used. This test consists in placing each test piece using different loads so that the lapping pressure will be 1, 2, 3, 4 and 5 kg/cm². It must be taken into account that the test will be done by using a 5 µm white alumina compound and grease as vehicle. This compound will be continuously fed between the lap and the workpiece during the lapping operation. Plots of stock removal rate (Q) versus the pressure (P) will be prepared.

On the other hand, the effect of abrasive size on the stock removal rate will be studied by using five different mesh size white alumina compounds. Likewise, a ½ inch diameter probe will be lapped at a fix lapping speed and pressure. Also, plots of stock removal rate (Q) versus the mesh size will be prepared.

Conclusions

The prototype design of a Rotating Lapping Machine with a Pneumatic Bearing has been completed and about 60% of the parts have been constructed. Tests are being planned to compare flatness and roughness of lapped surfaces produced by this machine with those produced by a conventional machine. Also the effects of lapping pressure and abrasive size on the stock removal rates will be investigated.

References

- [1] Lynah, P., Metals Handbook, vol. 3
- [2] *Grinding and lapping compunds*, <http://www.mfginfo.com/mfg/lapvalve/lapartical.htm>, United States Products, Co., Pittsburgh, PA, 1995.
- [3] Resendiz, R., et-al, *Chem.Eng.Tech.* **14** (1991), p.p. 105-108.
- [4] Kackzmarek, *Principles of machining by cutting, abrassion and erosion*, Peter Pergrinus, Lt., Warsaw, Poland.
- [5] Matsunaga M., *Fundamental Studies on Lapping*, Report of the Institute of Industrial Science, University of Tokyo, Japan, 1966.

SURFACE CONTAMINATION OF ABRASIVE DURING LAPPING

Adam W. BARYLSKI

Technical University of Gdańsk, Faculty of Mechanical Engineering
G. Narutowicza 11/12, 80-952 Gdańsk, Poland

1. Introduction

The high demands required today by design engineers for machine parts and tools necessitates very precise machining. The lapping process leads to a surface with low roughness and high precision. The topographical structure of the surface after lapping is very advantageous for sliding joints because of the high lubricant retention ability, as well as in non-sliding joints because of the high load - carrying ability. The range of lapped materials is very wide. This situation requires the use of natural and artificial abrasive materials and mainly such materials as micrograins of alumina, silicon carbide, boron carbide and diamond. Solid and liquid carriers distribute the micrograins on the surface of the lapping tool and the chemical active components make the machining more intensive [1].

One of the main disadvantages of lapping is the contamination of machined surfaces by micrograins of the abrasive material. The micrograins embedded into the machine surface during coarse lapping enter the „machining” zone during fine lapping. During service, contaminated surfaces are objected to excessive wear [3,5].

Since surface contamination is harmful and the methods of removal of the embedded micrograins are ineffective, the intensity of the contamination should be reduced as much as possible during the lapping process. The indentation of abrasive material into a machined surface should be considered as an element of general behaviour of abrasive micrograins including position of the micrograins, their size, shape as well as the physical and mechanical properties of the machined material and the lapping tool. The pressure and the lapping speed are also important.

2. Test Conditions

In the present investigation, surfaces contaminated by abrasive micrograins have been studied with an electron microscope and a microanalyzer. The combination of microscopy with X-ray analysis permits the study of secondary electron images and the X-ray images of the same spot on the sample [2,4,5]. The statistical size of abrasive slurry micrograins was estimated by means of an automatic image analysis (TV image analysis). The workpiece samples were made of spheroidal cast iron Zs55003 (2.6% C, 2.59% Si, 0.51% Mn, 0.15% P, 0.034% S, 0.15% Cu, hardness 1903 - 1942 MPa).

The samples were lapped in a special device with a planetary kinematic system. The paths of machined elements were of epicycloidal shape with almost constant speed of sample with respect to the lapping tool. The lapping disks were made of ferritic spheroidal cast iron Zs35022 (3.5% C, 2.3% Si, 0.51% Mn, 0.10% P, 0.045% S, 0.15% Cu, hardness 1491 - 1521 MPa).

3. Results

Tests were done in stages. In the first stage, the X-ray and spot analyses were carried out for the evaluation of indented micrograins with respect to the structural components of cast iron. In the second stage, the influence of lapping parameters such as pressure p , speed v and time t , on the contamination intensity were found. Complementary tests included the evaluation of the influence of micrograins size and concentration in the abrasive slurry as well as the method of cleaning the surface following lapping. Changes of radiation line intensity of specific elements in the X-ray analysis are equivalent to the changes of their concentration (Fig. 1).

Changes of radiation line intensity of specific elements in the X-ray analysis are equivalent to the changes of their concentration. At the same time the characteristic radiation intensity $\text{SiK}\alpha$ and $\text{FeK}\alpha$ were recorded at a sampling interval of $5\text{ }\mu\text{m}$. The pulse counting time was 10 s. The results of the contaminated surface analyse were related to the radiation value of non-contaminated surfaces of similar roughness.

Due to the fact that in the spot method the relative error depends on the number of analysed spots, the subsequent tests were based on the surface distribution analysis of the elements. Statistical test have proved that the distribution of number of indented abrasive micrograins observed in particular fields of analysis follows a Poisson distribution [1]. After estimating surface distribution for silicon it has been found that the concentration of indented abrasive micrograins is smaller for longer machining times.

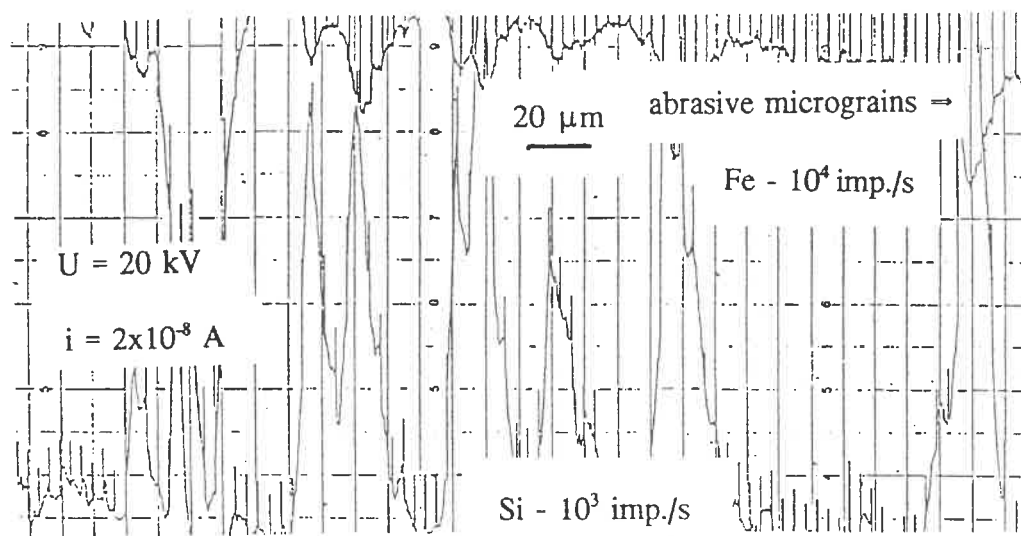


Fig.1. The linear distribution of Si and Fe on the surface of Zs55003 cast iron

Tests of indented micrograins have shown log-normal distribution for the micrograin size. Depending on the tested factors the mean value $E(d)$ of the indented micrograins size with respect to expected values of abrasive slurry micrograins amount to 30-60%. As well as surface analysis for the distribution of elements, secondary electron images have also been made. These observations show that surface contamination with micrograins of size $E(d) > 30\text{ }\mu\text{m}$. is very rare. The fragments of micrograins, a result of their cracking and chipping, penetrate into pearlite and ferrite as well (Fig.2). An increase in penetration or micrograins into the graphite border and into the ferrite envelope has been observed. The microphotographic analysis also shows that the micrograin shape is subjected to changes approaching isometric form where the rounding radius at the corners and the aspects angles are increased. The

influence of the packing factor B of the micrograin in the machining zone (Fig.3) and the micrograin size of cast iron contamination intensity have also been analysed [1].

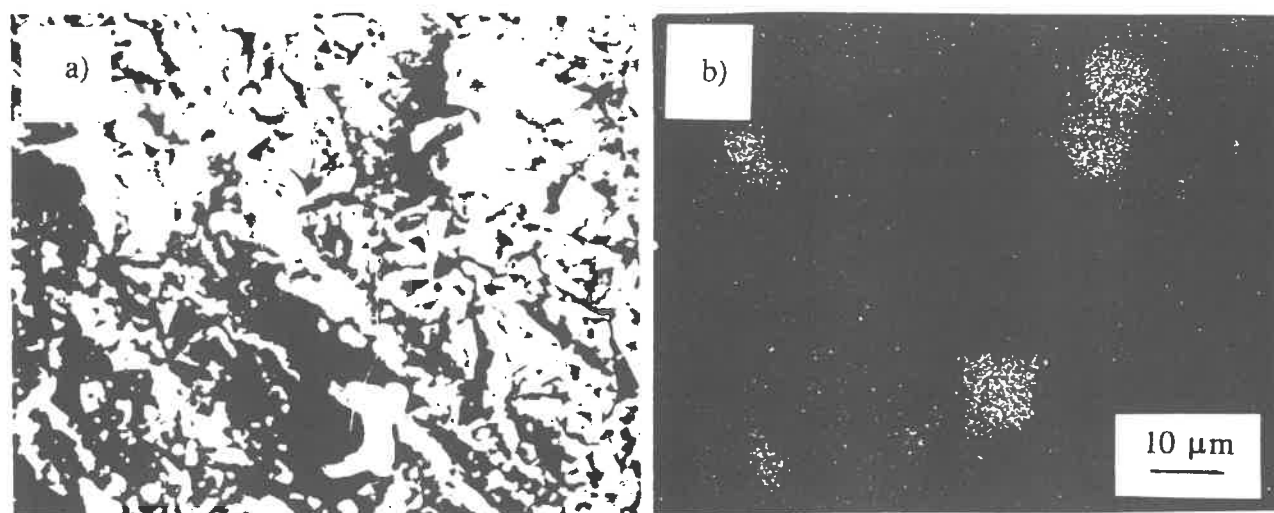


Fig.2. Microphotograph of Zs55003 cast iron: a) SEM, b) SiK α ($U = 25$ kV, JSM-35C)

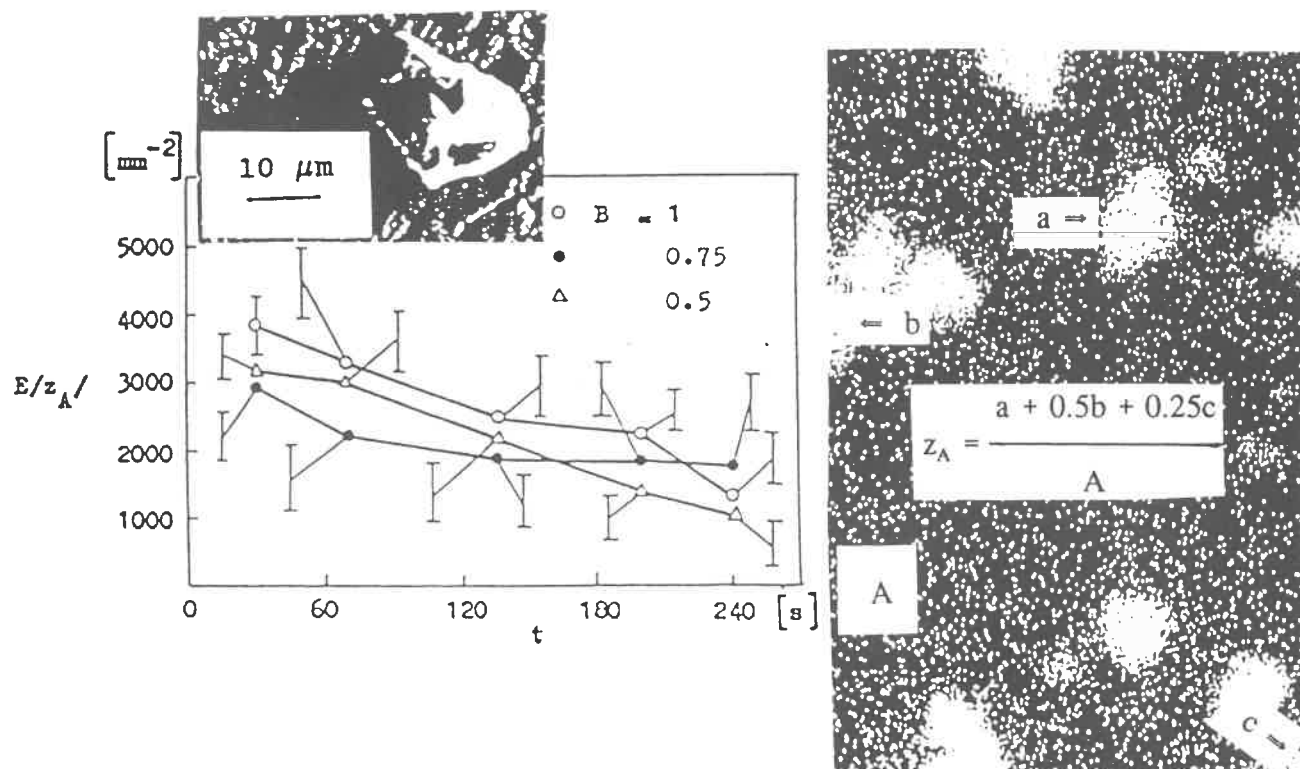


Fig.3. The effect of abrasive material packing factor B and the lapping time t on the concentration of micrograins contamination the surface of Zs55003 cast iron (silicon carbide 800, $p=0.19$ MPa, $v=0.46$ m/s)

4. Conclusions

The following conclusions can be formulated on the basis of the test that have been carried out:

- The effects of surface contamination with abrasive material can be reduced by prolonged time of lapping and by decreasing the micrograin packing in the machining zone in the beginning, or by reduction of unit pressure and lapping speed. During the lapping process the dispersion of abrasive micrograins increases whereas the load by a single micrograin and its penetration depth into the machined material is reduced.
- The distribution of size for silicon carbide micrograins indented into the machined surface is log-normal. The same type of distribution is typical for the characteristic radiation value $SiK\alpha$ defined in the cast iron surface after lapping.
- For greater micrograin number in the abrasive slurry, the concentration of cast iron contamination micrograins is reduced.
- There are no significant differences in intensity of surface contamination by silicon carbide micrograins after lapping and cleaning (manual or ultrasonic).
- The number of micrograins indented into cast iron surface per individual microscopic field of view has a Poisson distribution with relatively great variance, whereas the number of micrograins of defined size are characterized by a Poisson distribution of relatively small variance.
- Contamination of any particular structural components of the cast iron is relatively uniform [1].

References

1. Barylski, A.: Sc. Bull. of Gdańsk Tech. Univ., 491 (1992), 3-196
2. Barylski, A.: Pract. Met., 25 (1988), 349-355
3. Misra, A.; Finnie, I.: Wear, 65 (1981), 359-373
4. Read, S.J.: Electron Microprobe Analysis, Cambridge Univ. Press, 1975
5. Ronen, A., Malkin, S.: Wear, 68 (1981), 371-389
6. Takowitz, H.: Practical Scanning Microscopy, Plenum Press, New York, 1975

MODEL-BASED PRECISION MACHINING SIMULATION FOR THE PREDICTION OF TOOL LIFE

Katsuhiro Maekawa

Department of Mechanical Engineering, Ibaraki University
4-12-1 Nakanarusawa, Hitachi 316, Japan

INTRODUCTION

Tool life is of particular importance in metal machining since it influences machining economics and product quality: in particular, flank wear or a change in the shape of the edge directly leads to a geometrical error in the finished machining of a product. Taylor's tool life equation has been extensively used to estimate the cutting performance of a tool; however this requires a long series of wear tests entailing the extensive examination of variables such as cutting speed, feed, tool geometry and work material, which is time-consuming, wasteful of materials and expensive. A comprehensive, scientifically-based approach should be adopted to overcome these disadvantages; it is desirable to predict a tool-life diagram based on a wear characteristic equation in conjunction with a numerical simulation of the machining process.

The present paper describes a model-based machining simulation system for the prediction of tool life. Briefly discussing the system and the wear characteristic equation to be used, the way in which the wear of a P20 carbide tool develops with cutting time when a 0.46%C steel is machined in air at practical cutting speeds is demonstrated. The simulation results are then compared with experiments.

A MODEL-BASED MACHINING SIMULATION SYSTEM

Figure 1 shows the overall structure of a model-based machining simulation system for the prediction of tool life. The purpose of the concept is to obtain the maximum possible output of information on a given machining process using the smallest amount of experimental data. Initially required are finite element assemblage, cutting parameters and material properties. Neither contact stress nor tool temperature is prerequisite for the prediction of tool wear. A time-incremental analysis of the machining process based on a finite element analysis yields not only the progress of tool edge wear but also the changes in cutting forces, tool temperature, stress and strain distributions and chip morphology.

The core of the analytical method comprises finite element elastic-plastic and thermal analyses. In principle the following elastic-plastic constitutive equation and energy equation are solved simultaneously:

$$\dot{\sigma}_{ij} = \left[\frac{E}{1+\nu} \left(\delta_{ik}\delta_{jl} + \frac{\nu}{1-2\nu} \delta_{ij}\delta_{kl} \right) - \alpha^* \frac{9G^2}{\bar{\sigma}^2(3G+H')} \sigma'_{ij}\sigma'_{ij} \right] \dot{\epsilon}_{kl} \quad (1)$$

$$\dot{\theta} + v_j \theta_{,j} - K \theta_{,kk} = Q/\rho C \quad (2)$$

Equation (1) is the tensor expression for isotropic hardening materials obeying the Mises yield criterion and the Prandtl-Reuss flow rule, where $\dot{\sigma}_{ij}$ is the Jaumann rate of Euler stress, $\dot{\epsilon}_{kl}$ is the velocity strain, σ'_{ij} is the deviatoric component of Euler stress and $\bar{\sigma}$ is the equivalent stress; E , G , H' and ν are Young's modulus, the shear modulus, the work hardening rate and Poisson's ratio, respectively; δ_{ij} is the Kronecker delta; $\alpha^* = 0$ for elastic deformation and $\alpha^* = 1$ for the plastic state. In Eq. (2) θ is the temperature, $\dot{\theta}$ is the time derivative of temperature, v_j is the

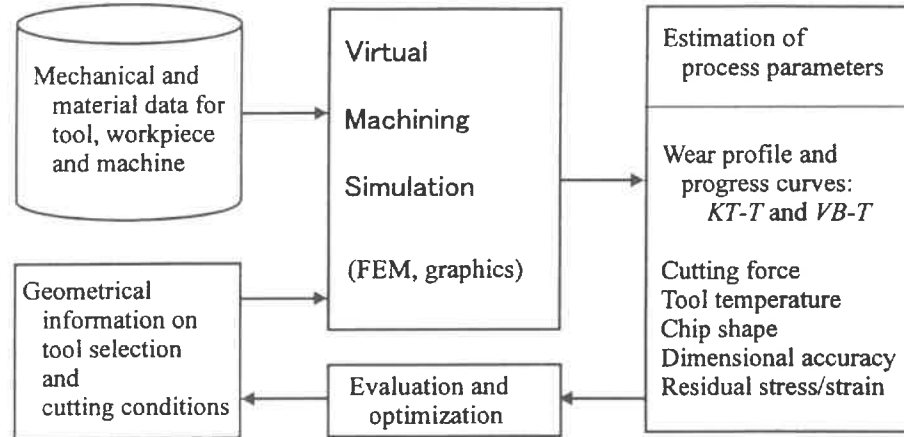


Fig. 1 Model-based machining simulation system for tool life prediction.

velocity, K is the thermal conductivity, Q is the internal heat generation, ρ is the density and C is the specific heat.

The finite element method can be employed to solve Eqs. (1) and (2) under the appropriate cutting geometry, material property and friction boundary conditions. The analysis applies to two-dimensional steady-state machining for which the iterative convergence method (ICM) is proposed. A review can be seen elsewhere [1, 2]. Since the tool is not rigid during machining its shape must be changed as both crater and flank wear develop. This can be done using an incremental method. When wear rate is estimated using a wear equation (see Eq.(5)) together with the temperature and normal stress obtained it is assumed that the tool wears away at the wear rate during a specified time increment, leading to slight changes in the shape of the edge. The ICM is repeated for the modified tool shape and information on chip geometry and flow stress within the chip is utilised as initial values at the subsequent interval. A steady state is also assumed at every time step so that the time does not correspond to the real machining time.

Good material property data are essential to the analysis. The uniaxial flow stress $\bar{\sigma}$ (MPa) of the workpiece is empirically expressed as a function of the strain $\bar{\epsilon}$, strain rate $\dot{\bar{\epsilon}}$ and temperature θ ($^{\circ}\text{C}$) according to

$$\bar{\sigma} = A(10^{-3}\dot{\bar{\epsilon}})^M e^{k\theta} (10^{-3}\bar{\epsilon})^m \left[\int_{\theta, \dot{\bar{\epsilon}} \equiv h(\bar{\epsilon})} e^{-k\theta/N} (10^{-3}\dot{\bar{\epsilon}})^{-m/N} d\bar{\epsilon} \right]^N \quad (3)$$

where the coefficients A , M and N vary with temperature and k and m are constants. These values are obtained from Hopkinson bar tests [3]. Friction conditions at the chip-tool and workpiece-tool interfaces are expressed as

$$\tau_t = k[1 - \exp\{-(\mu\sigma_t/k)^n\}]^{1/n} \quad (4)$$

where τ_t is the friction stress, σ_t is the normal contact stress, k is the shear flow stress and μ and n are constants. The coefficients are obtained from split tool tests [4].

Tool wear characteristics are a key factor in the present study, being formulated as

$$\frac{1}{\sigma_t} \frac{dW}{dL} = C_1 \exp\left(\frac{-C_2}{\theta}\right) + C'_1 \exp\left(\frac{-C'_2}{\theta}\right) \quad (5)$$

Equation (5) combines a thermally-activated adhesion model with a temperature-dependent abra-

sion model [5] and describes both crater and flank wear, where dW/dL is the wear rate per unit area, θ is the tool temperature (K) at the interface and C_1 , C_2 , C'_1 and C'_2 are constants to be determined for a combination of tool and work materials.

PREDICTED WEAR OF A CARBIDE TOOL IN STEEL MACHINING

The following case is typical: a 0.46%C plain carbon steel is orthogonally machined with a P20 carbide tool in air at a feed of 0.2 mm and cutting speeds of 100 and 200 m/min. The tool has a rake angle of 0° and a clearance angle of 4° . Table 1 lists the material properties measured and used in the simulation, including the values of the constants in Eqs. (3)–(5). The simulation programme installed on a UNIX workstation was interactively employed; it takes less than 10 minutes to obtain a wear progress curve together with a graphic presentation of other quantities such as chip shape and temperature distribution, as shown in Fig. 2.

Figure 2 shows a simulation result at a cutting speed of 100 m/min for (a) a sharp tool and (b) a worn tool when flank wear reaches a length of 0.6 mm. Due to the development of crater wear on the rake face a more curled chip is obtained. There is a marked difference in temperature distribution around the flank of the edge whereas no substantial changes can be seen on the rake face. The flank temperature rises to 700°C when machining the steel with the worn tool. Plots of the depth of the crater and the width of the flank wear land at each time-increment yield the wear progress curves shown in Fig. 3.

Table 1 Mechanical and thermal constants used in simulation

Flow stress: $A=1350 \exp(-0.0011\theta)+120 \exp\{-0.00006(\theta-275)^2\}$, $M=0.036$, $N=0.17 \exp(-0.001\theta) + 0.09 \exp\{-0.000015(\theta-340)^2\}$, $k=0.00014$, $m=-0.0024$ Young's modulus $E=206 \text{ GPa}$; Poisson's ratio $\nu=0.3$; Friction: $\mu=1.6$, $n=1.0$; Tool wear: $C_1=11.98 \text{ MPa}^{-1}$, $C_2=21950 \text{ K}$, $C'_1=7.80 \times 10^{-6} \text{ MPa}^{-1}$, $C'_2=5300 \text{ K}$			
	Thermal conductivity K $\text{Wm}^{-1}\text{K}^{-1}$	Density ρ kg m^{-3}	Specific heat C $\text{J kg}^{-1}\text{K}^{-1}$
Workpiece (0.46%C steel)	46.1	7840	502
Insert (carbide P20)	66.9	11200	356
Shank (0.55%C steel)	36.0	7750	461

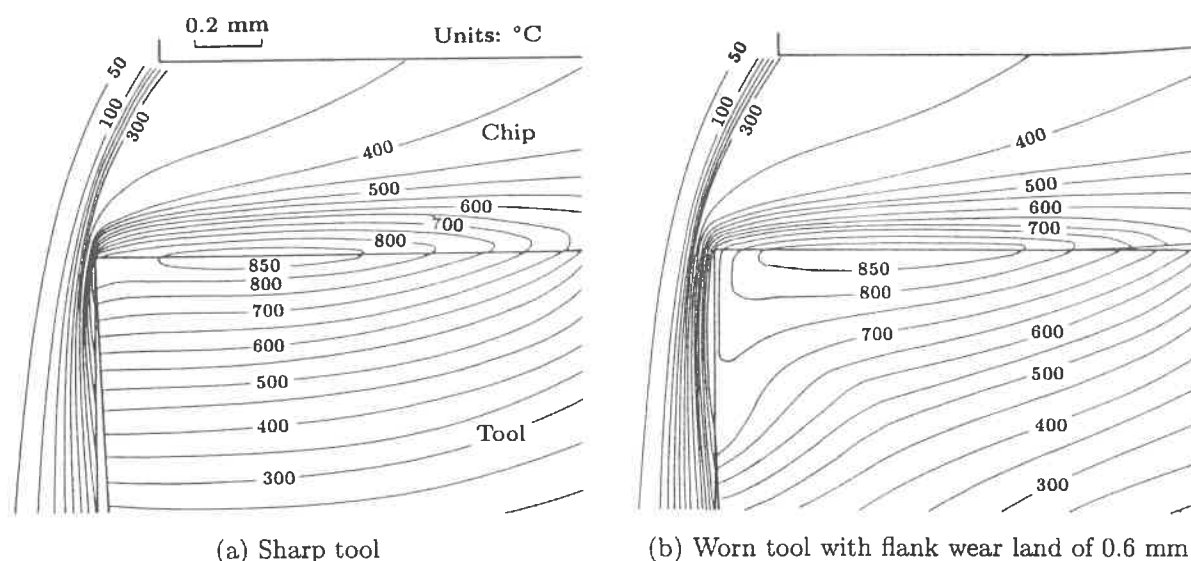


Fig. 2 Comparison of predicted chip shape and cutting temperature when machining 0.46%C steel with P20 tool at feed of 0.2 mm and cutting speed of 100 m/min.

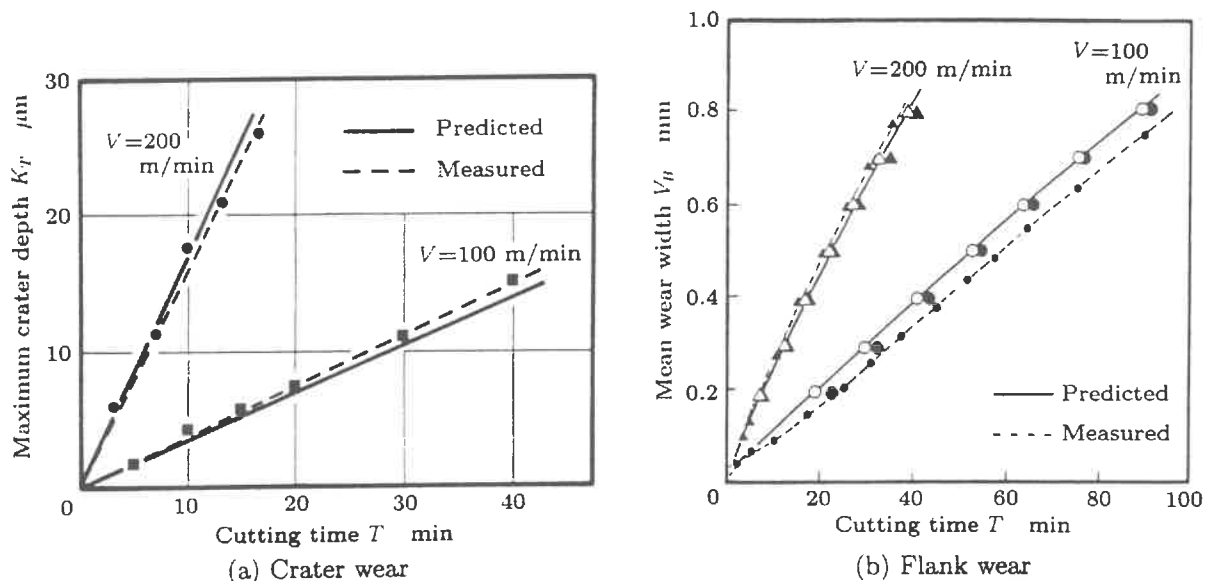


Fig. 3 Comparison of predicted tool wear progress curves with experiment: cutting conditions are the same as those in Fig. 2.

Figure 3 shows variations in tool wear as a function of time: (a) maximum crater depth and (b) mean flank wear width at cutting speeds of 100 and 200 m/min. The solid lines show the prediction and the dashed lines the experiment. Cutting experiments were conducted in the semi-orthogonal turning of a steel bar of $\phi 150\text{--}200 \times 300$ mm. The tool geometry was set at $(0^\circ, 0^\circ, 6^\circ, 6^\circ, 15^\circ, 15^\circ, 0.5\text{mm})$, tool shank dimensions were 30×30 mm and its overhang was 30 mm, which allows a stable cutting state without tool vibration. Reasonable agreement between the prediction and the experiment can be seen although an initial flank wear width of 0.03 mm is postulated in the simulation.

CONCLUSIONS

The present paper proposes a numerical method for the prediction of tool life in precision machining, in which an existing finite element model [2] is extended to cope with the deterioration of a cutting edge. A numerical simulation based on a wear characteristic equation is carried out to show how the wear of a cutting tool is developed with increasing cutting time. As a result tool life can easily be obtained in conjunction with changes in other quantities including cutting force, tool temperature and stress and strain in the work material. The applicability of the proposed method is shown by comparison with experiments when machining a 0.46% C steel with a P20 carbide tool under controlled laboratory conditions.

Acknowledgments – The author would like to thank Professor T. Kitagawa, Department of Mechanical System Engineering at Kitami Institute of Technology, Japan, for his valuable comments.

REFERENCES

- [1] Usui, E. and Shirakashi, T., On the art of cutting metals – 75 years later, ASME Publ. PED-7 (1982) 13–35.
- [2] Maekawa, K., Shirakashi, T. and Obikawa, T., Proc. Instn Mech. Engrs, Journal of Engineering Manufacture, B 210 (1996) 233–242.
- [3] Shirakashi, T., Maekawa, K. and Usui, E., Bull. Japan Soc. Prec. Engng, 17 (3) (1983) 161–172.
- [4] Maekawa, K., Kitagawa, T. and Childs, T.H.C., Proc. 2nd Int. Conf. on Behaviour of Materials in Machining, York, UK, Nov. 14–15 (1991) 132–145.
- [5] Kitagawa, T., Kubo, A. and Maekawa, K., Proc. 23rd Leeds-Lyon Symposium on Tribology, Leeds, UK, Sept. 10–13 (1996).

AN ORTHOGONAL CUTTING STABILITY MODEL WITH VARIABLE DELAY LENGTH DUE TO TOOL MOTION

Thomas B. O'Reilly, Eric R. Marsh
The Pennsylvania State University
University Park, PA

Background

Regenerative chatter in machine tools occurs when irregularities in the surface of the part being machined grow with each pass of the tool. This is a severe problem in manufacturing industries which results in poor surface finish, poor dimensional accuracy and increased tool wear. Once chatter marks have developed in a workpiece they are very difficult to remove due to the regenerative effect which occurs as the tool cuts the material. The chatter marks left by previous passes of the tool cause the structure to vibrate in such a way that chatter marks of larger amplitude are produced with each pass. In some cases it is faster and easier to scrap the part and start over resulting in lost time and material.

Complicating the stability analysis for most types of machining processes is the fact that the forcing function that the tool experiences is related to both the current position of the tool and the profile of the surface that is being removed. This in turn is related to the position of the tool the last time it passed over that part of the surface. As a result, time delay terms appear in the equations of motion.

Figure 1 shows a typical model used to predict the onset of chatter in orthogonal, or plunge cutting, as would be performed on a lathe. This model only accounts for motion of the tool into and out of the workpiece. Most existing models used to predict the chatter limit, such as the one presented in Tlustý [4], are based on this type of analysis. For a model of this type with constant spindle speed the length of the delay term in the equations of motion is constant.

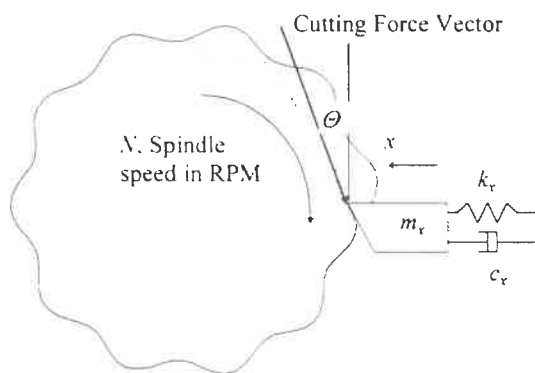


Figure 1. Diagram of a workpiece and cutting tool on a single degree-of-freedom spring/damper system.

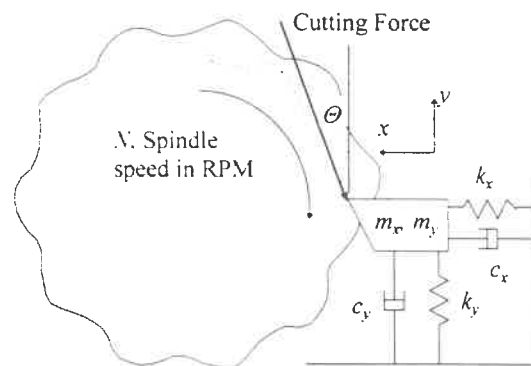


Figure 2. Diagram of a workpiece and cutting tool on two degree-of-freedom spring/damper system.

As part of research sponsored by the National Institute of Standards and Technology a new model of chatter has been developed which accounts for the motion of the tool in both the x - and y -directions as shown in Figure 2.

Analytical Model

The system was first examined using an analytical model. The two degree-of-freedom model used in this study assumes that the cutting force proportional to the area of the chip, and neglects the steady-state offset induced by the mean cutting force. In addition the stiffness and damping are assumed to be linear.

The equations of motion are:

$$X\text{-direction: } m_r \ddot{x} + c_r \dot{x} + k_r x = -\alpha F \quad \text{where: } \alpha = \sin \Theta \quad (1)$$

$$Y\text{-Direction: } m_r \ddot{y} + c_r \dot{y} + k_r y = -\beta F \quad \text{where: } \beta = \cos \Theta \quad (2)$$

Cutting Force Approximation:

$$F = bC(x(t) - x(t - \tau)) \quad (3)$$

Where b is the width of the cut, C is a material dependent cutting stiffness term, and τ is the length of time since the tool last passed over this point on the workpiece. For the single degree-of-freedom case this amount is fixed and is given by:

$$\tau = 60 / N \quad \text{where } N \text{ is the spindle speed in RPM.}$$

For the two degree-of-freedom model the amount of the delay is not constant due to the motion in the y -direction. Because constant delay terms are much easier to deal with, it is necessary to convert the equations of motions to a form in terms of arc length, u , defined as follows:

$$u = Vt + y$$

where V is the linear velocity of the tool relative to the workpiece, and t is time. After performing this transformation the equations of motion can be written as:

X -direction:

$$m_r \left(\frac{du}{dt} \right)^2 x''(u) + c_r \frac{du}{dt} x'(u) + k_r x(u) = -\alpha b C(x(u) - x(u - d\pi)) \quad (4)$$

Y -direction:

$$m_r \left(\frac{du}{dt} \right)^2 y''(u) + c_r \frac{du}{dt} y'(u) + k_r y(u) = -\beta b C(x(u) - x(u - d\pi)) \quad (5)$$

Terms such as x' , y' indicate derivatives with respect to arc length s , rather than time. The delay term is now the constant length $d\pi$, where d is the workpiece diameter, rather than a varying length of time. In these equations:

$$\frac{du}{dt} = V + \dot{y} \approx V$$

Since $\dot{y}$ should be small compared to V the system is linearized by neglecting the $\dot{y}$ term.

The tau-decomposition method (presented in Hsu [1], Lee [2] and Stepan [3]) was then used to examine the stability of this system and produce the stability chart shown in Figure 3. On this plot the straight horizontal line is based on the results of Tlustý [4], who derived the stability limit for a single degree-of-freedom model. The lower lobed line was produced using the tau-decomposition method and a single degree-of-freedom model of the system. The upper lobed line was produced by the two degree-of-freedom model described above. In all three cases the system parameters in the x-direction are the same.

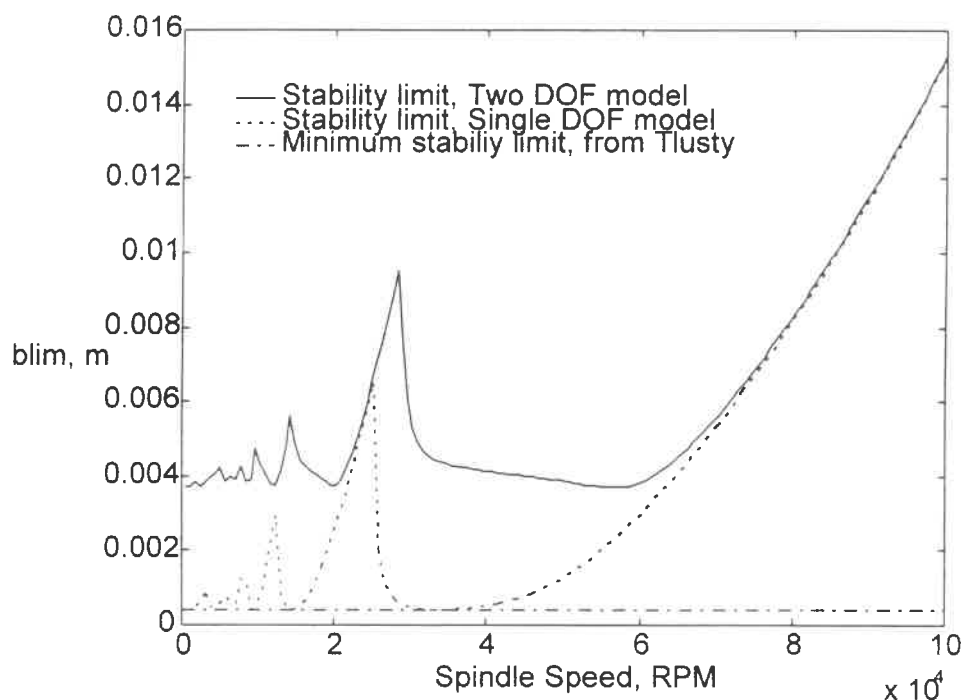


Figure 3. Two degree-of-freedom stability chart.

Numerical Simulation

A numerical simulation of this system was written to examine the dynamics of the cutting process under a variety of conditions. Using this simulation, the effect of different ratios between the x- and y-direction stiffness and natural frequencies was studied to determine its impact on the stability of the system and on the types of tool motions seen during cutting. The

simulatic. can also be used to numerically generate stability charts for comparison to the simplified analytical model and experimental data.

Experimental Work

A two degree-of-freedom flexure was designed and constructed to experimentally test this model. The stiffness of the flexure is much lower than the stiffness of the machine tool base. Therefore the dominant dynamics of a tool mounted in the flexure are due to the compliance of the flexure. In this way the behavior of the cutting tool approximates the model used to develop the stability predictions above.

Orthogonal cuts of different widths were made using a steel tool mounted in this flexure. Cutting tests were performed under a variety of conditions to study the effect that varying speeds and feed rates have on the system.

Conclusions

The initial results obtained using the model of chatter described above suggest a large gain in stability over a single degree-of-freedom model. Although this effect has not been verified experimentally, experimental data shows some effects that are worth investigating, some of which may be due to the y -direction motion. It is possible that the simplifications made in the system model are not appropriate, in which case the model may be overestimating the stability of the actual process. A more complete examination of the system, including the nonlinearities related to the delay term, is being performed using numerical models. In addition, more detail may be included such as a more comprehensive model of the cutting force, and the effect of the curvature of the workpiece.

To verify this more comprehensive model an experimental stability chart will be produced using the flexure. These results will also be compared to existing experimental stability charts found in the literature.

References

1. Hsu, C. S. "Application of the Tau-Decomposition Method to Dynamical Systems Subjected to Retarded Follower Forces," Journal of Applied Mechanics, June 1970, pp 259-266.
2. Lee, M. S., and C. S. Hsu. "On the τ -Decomposition Method of Stability Analysis for Retarded Dynamical Systems," SIAM Journal on Control, Volume 7, 1969, pp 242-259.
3. Stepan, G. Retarded dynamical systems: stability and characteristic functions, Longman Scientific & Technical, Essex, UK, 1989.
4. Tlustý, J. Machine Dynamics (1985). In King, Robert I. (ed), Handbook of High Speed Machining Technology (pp. 48-143). New York, Chapman and Hall.

FUZZY LOGIC MODELING OF ELECTROLYTIC GRINDING PROCESS

K. G. R. Koteswara Rao, G. Kuppuswamy
Indian Institute of Technology
Madras, INDIA

Key words : Electrolytic Grinding, Machinability Assessment, Fuzzy logic, Modeling and Simulation

Introduction

Electrolytic Grinding, one of the precision processes, is a material removal process by both mechanical and electrolytic actions. The process employs an electrically conductive metal bonded wheel on the anodic workpiece with an electrolyte introduced in the electrode gap. Due to numerous interacting variables involved, the mathematical modeling is very difficult to be formulated for process analysis. The nature of such machining responses are largely probabilistic rather than deterministic [1].

Need for fuzzy logic approach

Fuzzy logic modeling is based on mathematical theory combining multivalued logic, probability theory and Artificial Intelligence methods and can be used to tackle complex problems. The main advantage of Fuzzy logic analysis is that there is no need to establish a mathematical model. Fuzzy modeling is based on fuzzy set theory and the linguistic statements expressed mathematically which corresponds to the analysis of a human expert. Fuzzy systems base their decisions on inputs and outputs in the form of linguistic variables. The variables are tested with IF-THEN rules, which produces one or more responses depending on which rules are asserted. The response of each rule is weighed according to the degree of membership of it's inputs and the centroid of the responses is calculated to generate the appropriate output [2&3]. Typically the input variables for the face grinding process are feed force and current density, the output variables being material removal rate or grinding time and surface roughness.

Experimental details

The experimental set up shown in Fig. 1, consists of an electrolytic diamond wheel (type D6A2C) of diameter 150 mm run at 3000 rev/min. and 120 double oscillations per minute with a stroke of 6 mm. A 12V d.c. supply was used. The electrolyte was a solution of 15 % wt. NaNO_3 in water. WC tool bit (K20 grade) was face ground [4].

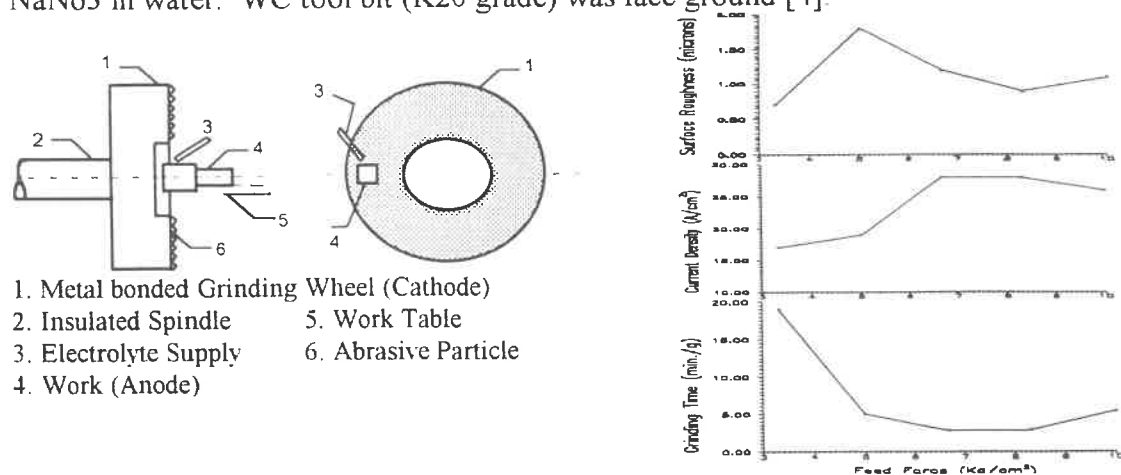


Fig 1: Electrolytic Grinding process and experimental results

Fuzzy logic modeling

The basic steps involved in the fuzzy logic analysis are as follows[5]:

1: Defining inputs and outputs

First, the ranges for the variables are identified, that represent the relevant conditions of the given process. The range of values that inputs and outputs may take is called the universe of discourse or dimension.

2: Fuzzify the inputs and outputs

Then meaningful linguistic statements are selected for each variable and expressed by appropriate fuzzy sets. The fuzzification for the feed force and current density are given in the Table 1 and Figs. 2 & 3 represent the fuzzy membership values for these variables.

Table 1: Fuzzification of the input variables

Feed force (kg/cm^2):

Fuzzy variable	Crisp Input Range
Very Low (VL)	0-4
Low (L)	2-6
Medium (M)	4-8
High (H)	6-10
Very High (VH)	8-12

Current Density (A/cm^2):

Fuzzy variable	Crisp Input Range
Low (L)	0-15
Medium (M)	7.5-22
High (H)	15-30

The change of states, like very low to low, is assumed to be continuous rather than abrupt and are represented by the set of overlapped triangles. Fuzzy memberships are assigned to the output variables just as for input variables and these are given in the Table 2 .

Table 2 : Fuzzification of the output variables

Grinding Time ($\text{min.}/\text{g}$):

Fuzzy variable	Crisp Input Range
Extremely Low (XL)	0-6
Very Low (VL)	3-9.5
Low (L)	6-12
Medium (M)	9.5-15
High (H)	12-18
Very High (VH)	15-22
Extremely High (XH)	18-25

Surface Roughness (microns):

Fuzzy variable	Crisp Input Range
Very Low (VL)	0-0.6
Low (L)	0.3-0.9
Medium (M)	0.6-1.1
High (H)	0.9-1.4
Very High (VH)	1.1-1.7
Extremely High (XH)	1.4-2.0

3: Create a fuzzy rule base

Table 3 refers to the fuzzy rule base matrix and these had been developed by heuristic guidelines of the analyst based on the experimental results, and represents the knowledge base.

Table 3 : Fuzzy rule matrix

F.F→ C.D↓	VL	L	M	H	VH
L	G.T:XL S.R:VL	G.T:H S.R:L	G.T:L S.R:M	G.T:XL S.R:H	G.T:L S.R:XH
M	G.T:VH S.R:VL	G.T:M S.R:L	G.T:VL S.R:M	G.T:XL S.R:VH	G.T:L S.R:XH
H	G.T:H S.R:L	G.T:M S.R:M	G.T:XL S.R:H	G.T:VL S.R:VH	G.T:L S.R:XH

F.F - Feed Force
C.D - Current Density
G.T - Grinding Time
S.R - Surface Roughness

4: Defuzzify the output

The inputs and outputs to the fuzzy rule base are fuzzy variables. The final step is defuzzification of the output for the given inputs into a crisp value. This is normally computed by either Fuzzy-OR method or centroid method of defuzzification on the fired rules.

Case Example:

For a feed force of 7.3 Kg/cm² removal and current density of 25 A/cm², the grinding time estimate is to be computed. By fuzzification, feed force comes under two categories, namely Medium and High fuzzy sets with 0.33 and 0.69 membership values respectively. Current Density comes under the High category with membership value 1.0.

Two rules are fired from the knowledge base rule matrix:

1) IF Feed force IS Medium (0.33) AND Current Density IS High (1.0)

THEN Grinding time IS eXtremely Low.

2) IF Feed force IS High (0.69) AND Current Density IS High (1.0)

THEN Grinding time IS Very Low.

First the firing strengths of the rules are determined from the two rules by the conjunction or minimum operator and these represent membership values in the corresponding fuzzy sets.

For rule 1) : $0.33 \wedge 1.0 = 0.33$ of category eXtremely Low (XL) of Grinding time

For rule 2) : $0.69 \wedge 1.0 = 0.69$ of category Very Low (VL) of Grinding time

Final output value of the Grinding time is obtained by the centroid method of defuzzification as follows.

The strength values or the fuzzy set memberships of the output after firing of the rules are used to fill in the concerned triangular membership functions of the output parameter to that strength level in the fuzzy sets diagram. So category Extremely Low (XL) is filled up to 0.33 and Very Low (VL) is filled up to 0.69 as shown in the Fig. 4. These truncated figures are aligned and the centroid of that overlapped figure gives the output value of 4.8 min./gm. Similarly surface roughness value for the given inputs is 1.3 microns as given in Fig. 5.

Figs. 6&7 represent the output estimate surface for grinding time and surface roughness respectively based on similar computations. Figs. 8&9 represent the two dimensional views having one input variable fixed. Thus the method becomes a simulation tool for further process analysis.

Conclusions

The fuzzy logic inference methodology for modeling of Electrolytic Grinding Process appears to be a more meaningful alternative to the traditionally known methods for predicting output estimates quantitatively. The methodology presented is suitable in principle to any arbitrary machining problem provided that the associated knowledge base could be incorporated appropriately.

References

- 1) Iwata K Murotsu Y and Fujii S. A Probabilistic Approach to the Determination of the Optimum Cutting Conditions. ASME J. of Engg. for Ind., 1972, November, 1099-1107.
- 2) Zadeh L.A. Fuzzy sets. Informat. Contr., 1965, vol. 8, 338-353.
- 3) Schwartz D.G., Klir G.J., Lewis III H.W. and Ezawa Y. Applications of Fuzzy sets and Approximate Reasoning. Proceedings of the IEEE, 1994, April, 482-498.
- 4) Kuppaswamy G. Wheel variables in electrolytic grinding. Tribol. Int., 1976, 9(1), 29-32.
- 5) Martin McNeill F., Thro E. 'Fuzzy Logic-A Practical Approach'. AP Professional, 1996.

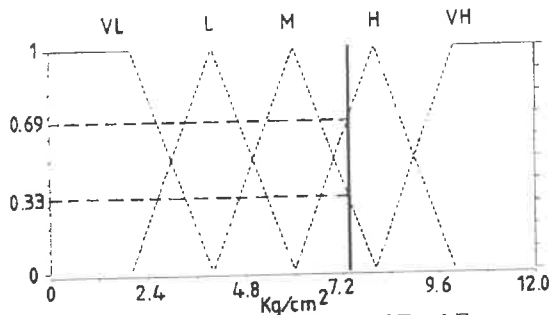


Fig 2 : Fuzzification of Feed Force

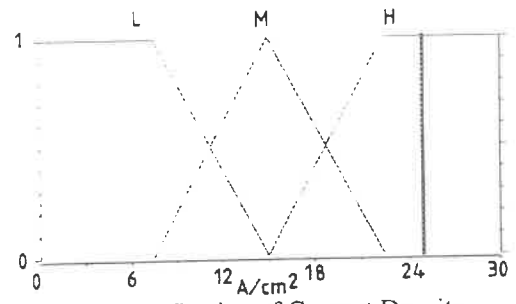


Fig 3 : Fuzzification of Current Density

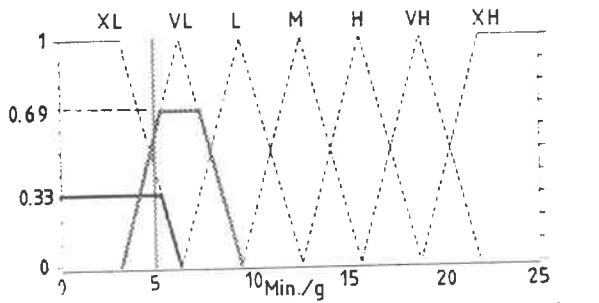


Fig 4 : Defuzzification of Grinding Time

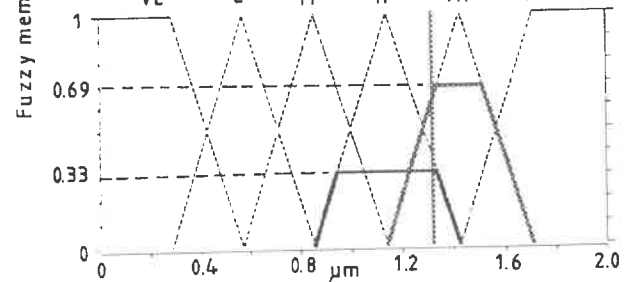


Fig 5 : Defuzzification of Surface Roughness

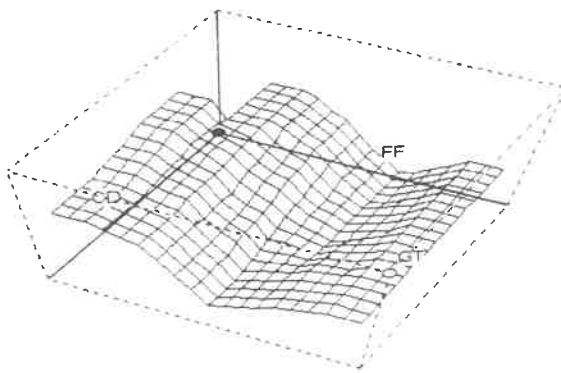


Fig 6 : 3-D View of the output against inputs
(For Grinding Time)

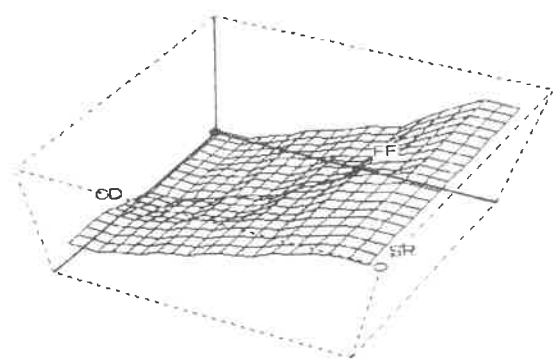


Fig 7 : 3-D View of the output against inputs
(For Surface Roughness)

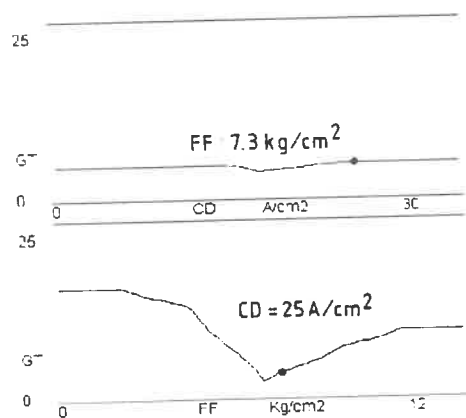


Fig 8 : 2-D View of the output against inputs

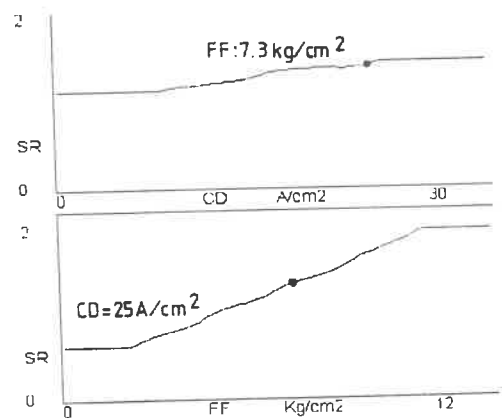


Fig 9 : 2-D View of the output against inputs

AN ADAPTIVE METHOD FOR PREDICTING RESIDUAL STRESSES ASSOCIATED WITH GRINDING INDUCED PHASE TRANSFORMATION

Mofid Mahdi and L. C. Zhang

Department of Mechanical and Mechatronic Engineering
The University of Sydney, NSW 2006, Australia

Abstract

This paper developed an efficient adaptive method to simulate the thermal residual stresses of grinding, particularly those associated with phase transformation when the workmaterial was subjected to a critical grinding temperature variation. The change of workpiece properties was modelled as a function of grinding temperature history to account for the surface hardness and volume increase during phase transformation. It found that if the properties of a workpiece is temperature-independent, the residual stresses in both the grinding direction and that perpendicular to it are tensile. The maximum residual stress in this case is about 1.25 times the initial yield stress of the material and is not affected by the further increase of the input heat flux generated by grinding. When the material properties are temperature-dependent, however, the increase of the surface hardness due to phase transformation will raise the level of the maximum residual stress to about 2.7 times the original yield stress. The volume growth associated with phase change plays an important role in the formation and nature of residual stresses.

NOTATIONS

Cr_{700}	the cooling rate at 700°C
H_v	Vickers hardness
l_a	relative peak location of a heat flux ($2\zeta_a/L_c$), see Fig 1
M	martensite
P_c	Pecklet number ($vL/2\alpha$)
q	heat flux per unit grinding width
q_a	peak value of the heat flux
T	temperature rise with respect to ambient temperature T_∞
v	moving speed of the heat source, see Fig 1
Y	yield stress of the workmaterial
α	thermal diffusivity
σ	stress tensor

Subscripts

x, y, z	x-, y-, and z-directions, see Fig. (1)
y	yield
∞	room temperature

1. INTRODUCTION

The rapid heating and cooling in a grinding process may cause phase transformation and thermal plastic deformation in a workpiece and in turn introduce substantial residual stresses. The present study aims to develop an efficient adaptive approach to investigate the grinding residual stresses of phase transformation with the aid of the finite element method taking into account the property changes of workpiece materials

due to phase transformation.

2. MODELLING

As shown in Fig. 1, the deformation of a workpiece subjected to a surface grinding operation is considered as a plane-strain problem. The heat generated by grinding is approximated by a triangular heat source moving along the positive direction of x-axis on the workpiece surface [1].

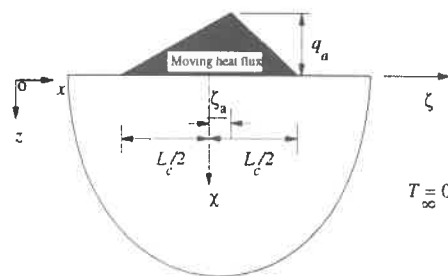


Fig. 1 Thermal grinding model

2.1 Material Properties

When a workmaterial experiences a critical temperature change in grinding, phase transformation will take place [1-2]. To study the residual stresses of phase transformation by the finite element method, We propose the following constitutive model for the EN23 steel alloy to describe the irreversible

deformation induced by martensite transformation. The variation of martensite hardness, H_v , can be determined by the steel composition directly [3-4], i.e.,

$$H_v = 463.2376 + 21.0 \log(Cr_{700}) \quad (1)$$

The yield stress of martensite, σ_y , is calculated by

$$\sigma_y = 80.0 + 2.7544 H_v \quad (2)$$

The yield stress of the workmaterial surface undergoing phase transformation is treated as a weighted average quantity based on the martensite percentage formed and original yield stress, Y . To demonstrate the grinding temperature history a typical transient surface grinding temperature is considered in relation to heat source movement (Fig. 2).

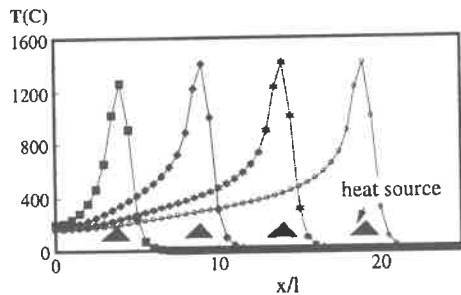


Fig. 2 Typical surface temperature ($q_a=80 \text{ MW/m}^2$, $l_a = 0.25$, $Pe_\infty = 1$)

The figure indicates that the surface temperature approaches the steady state in a short time. For the surface layer in the vicinity of the ground surface, ($z=0$), the progress of the martensite yield stress profile has a strong relation to heat source (see Fig. 3).

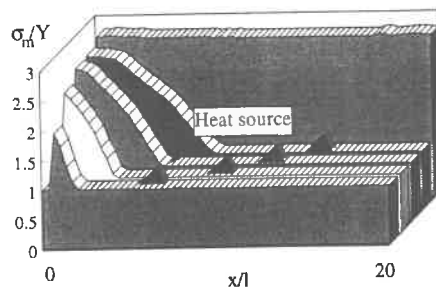


Fig. 3 Typical surface yield stress vs heat source movements ($q_a=80 \text{ MW/m}^2$, $l_a = 0.25$, $Pe_\infty = 1$)

In an early stage, the development of martensite phase creates a hardened zone. It is clear that the surface yield stress increases to a level of $2.8 Y$ as

more martensite forms.

According to the above understanding, an adaptive algorithm is built up, see Fig. 4, to account for the general change of material properties and perform the nonlinear calculations based on the grinding conditions,

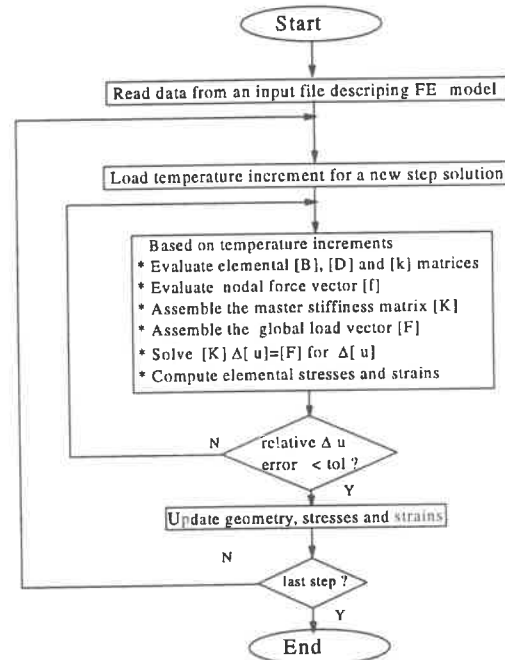


Fig. 4 An adaptive algorithm

When grinding conditions are not critical, the yield stress is constant. The results of developed algorithm are compared with those calculated by ADINA code [5] for the same grinding conditions and show close agreement in terms of the maximum residual stresses for different grinding conditions (Fig.5).

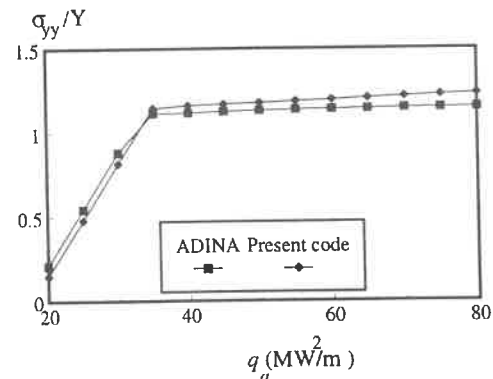


Fig. 5 Residual stresses (ADINA vs present code) ($q_a=80 \text{ MW/m}^2$, $l_a = 0.25$, $Pe_\infty = 1$)

3. RESULTS AND DISCUSSION

Figure 6 compares the variation of residual stresses with surface hardening. In the case of surface hardening, stresses σ_{xx} and σ_{yy} are both tensile with a maximum of 2.6 times the initial yield stress of workpiece material. Therefore the hardening effect results in a higher tensile stresses since the yield stress of martensite is higher. The variation of residual stresses near the surface is nearly linear for both stresses but with a rapid change at the onset of phase transformation. At a higher depth, the residual stresses are similar for both cases regardless of surface hardening. This indicates that the residual stresses within the martensite zone is directly related to martensite depth. On the other hand, the permanent volume growth of martensite is restrained by the part without transformation thus results in compressive residual stresses (Fig. 7). The effects of the surface hardening and the volume growth are combined and considered in Fig. (8). The combination results in higher magnitudes of residual stresses regardless of their nature.

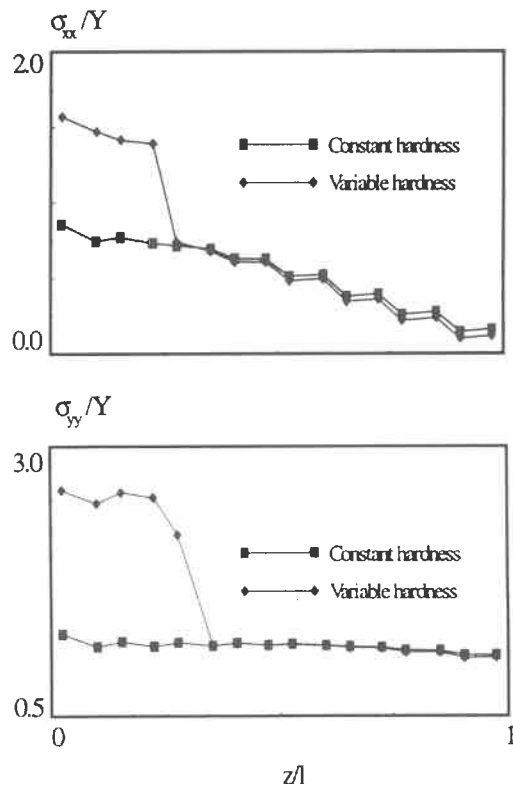


Fig. 6 Typical residual stress distribution vs effect of surface hardening ($q_a=80 \text{ MW/m}^2$, $l_a = 0.25$, $Pe_\infty = 1$)

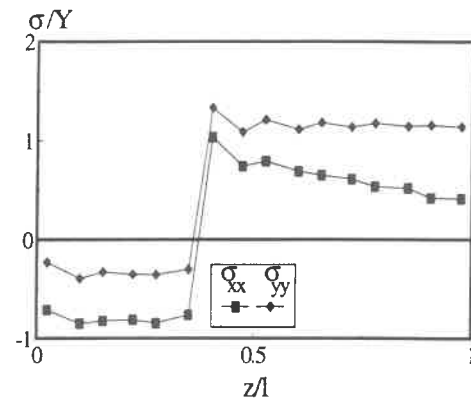


Fig. 7 Residual stress distribution vs effect of volume growth ($q_a=80 \text{ MW/m}^2$, $l_a = 0.25$, $Pe_\infty = 1$)

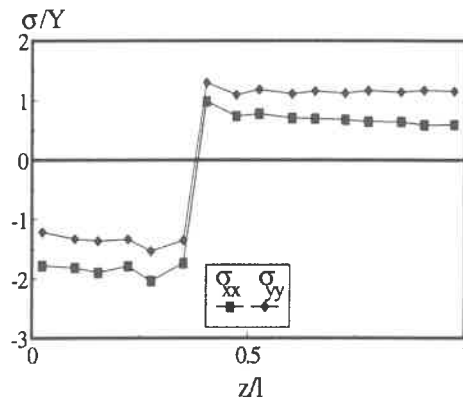


Fig. 8 Residual stress distribution vs effect of volume growth and surface hardening ($q_a=80 \text{ MW/m}^2$, $l_a = 0.25$, $Pe_\infty = 1$)

4. CONCLUSIONS

A mathematical material model and a feasible adaptive algorithm have been developed for studying the residual stresses of grinding induced by phase transformation. Analysis can be carried out at the resolution of integration points. It concludes that (1) the developed algorithm is a useful and stable tool, (2) the workmaterial composition is a key factor in determining the phase change and the associated constitutive relations, and (3) the volume growth and surface hardening dominate the transition of residual stresses from compressive to tensile.

Acknowledgment- The financial support from ARC Small Grant to the present study is appreciated. ADINA code was used for grinding temperature calculations.

REFERENCES

- 1 M. Mahdi and L. C. Zhang, The finite element thermal analysis of grinding processes by ADINA, *Computer & Structure*, **56**, (2/3) 313-320 (1995).
- 2 L. C. Zhang and M. Mahdi, Applied mechanics in grinding, part IV: grinding induced phase transformation. *Int. J. Mach. Tools Manufact.*, **35** 1397-1409 (1995).
- 3 British Iron and Steel Research Association, *The mechanical and physical properties of the british standard on steels*, Vol.2, Pergamon Press ,Oxford, UK (1969).
- 4 M Atkin BSc, B Met,FEM , *Atlas of Continuous Cooling Transformation diagrams for Engineering Steels*, Market Promotion Department, British Steel Corporation, BBC Billet, Bar and Rod Products, Sheffield, UK (1977).
- 5 ADINA R & D, inc, *Theory and modelling guide*. Report ARD 92-5, ADINA R & D, inc, December, (1992).

Transient Response of Wire Electrode in Wire Electrical Discharge Machining

Hisashi Yamada, Notake Mohri, Katsushi Furutani

Toyota Technological Institute, Nagoya 468, Japan

Takuji Magara

Mitsubishi Electric, Corp., Nagoya 461, Japan

1. Introduction

In wire electrical discharge machining (wire-EDM), fine machining with high accuracy can be realized since a fine thread wire is used as a tool electrode. Recently, there is a demand of higher accuracy in the manufacture of fine products. Thus, it is necessary to restrain the vibration of a wire caused by a force received in wire-EDM.

A numerous number of works have been done concerning the vibration of the wire electrode in steady state based on long term observation [1][2]. However, there are few papers discussing the transient behavior of a wire electrode forced by each discharge which moves discontinuously on the wire electrode. The authors investigated the randomness of wire displacement caused by concentrated force at a certain point during a thin plate

cutting. It is found that an averaged external force acts on the wire electrode as an exciting force or a restraint force according to the direction of the wire movement at the moment when discharges occur. This is the reason why the amplitude of the wire vibration changes randomly [3].

In the analysis of wire vibration in steady state, wire can be approximated as a string. However, the force caused by electrical discharge acts on the wire intermittently, very locally and longitudinally. Thus, in order to analyze the wire vibration in detail when the force caused by discharge acts on the wire, it is necessary that the transient phenomena are to be investigated.

In this paper, in the discussion of transient vibration of the wire, we represent the wire as a beam under a constant tensile force. The vibration process of the wire is considered using experimental results right after voltage is supplied between the wire electrode and the workpiece.

Table 1 Experimental conditions

Wire electrode	$\phi 0.200\text{mm}$, brass
Wire guide	Internal diameter: $\phi 0.205\text{mm}$
Work piece	Steel 1mm and 20mm in thickness
Pulse setting	Open-circuit voltage u_i : 80,150V Peak current of discharge i_e : 145A Discharge duration t_e : 0.7 μs
Distance between two support guides: L	80mm
Tension of wire: T	9.5N
Wire winding rate	145mm/s
Working fluid	Deionized water(Conductivity:15 μS) Flow rate: upper 1.0l/min, lower 0l/min

2. Wire vibration caused by single discharge

Before dealing with the vibration caused by continuous discharges, firstly it is necessary to investigate the response of wire electrode to a single discharge. Therefore, the transient response of the wire to a single discharge is observed. The measurement system is shown in Fig. 1 and the experimental conditions are shown in Table 1. The wire electrode is supported by an upper and a lower supporting guides, wound downward at a constant speed, and stretched by tensile force. It is bent by electrical contacts for supplying machining pulse voltage and assisting guides as shown in Fig. 1. Both of the supporting guides are made up of diamond having a hole with a clearance of 5mm, so that it enables the wire to pass through. The workpiece plate of stainless steel with 1mm in thickness is cramped at the height of the center between two supporting guides, so that the center is located approximately at the center of the workpiece in vertical direction.

The displacements of the wire electrode are measured with an optical displacement meter (OPT-FOLLOW 7100C, YA-MAN Corp., Frequency bandwidth: 500kHz). The measuring location is set up at M , so that the measured displacement is not disturbed by

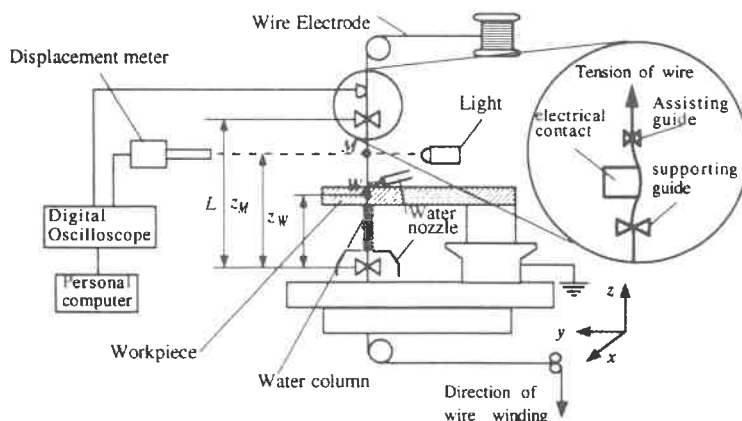


Fig. 1 Measurement system for wire vibration in wire-EDM

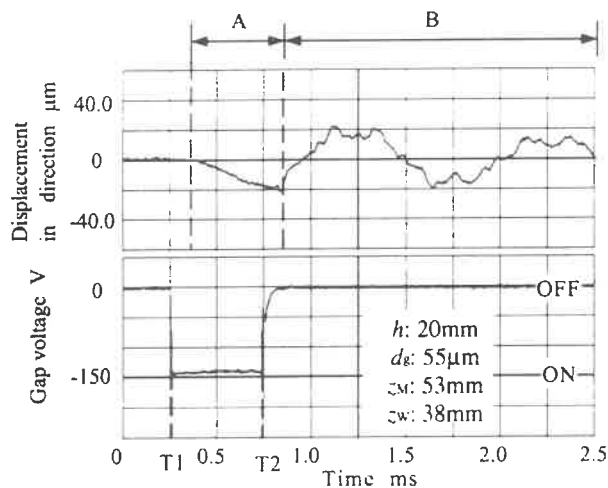


Fig. 2 Transient response of wire to supplied voltage and discharge

flushing deionized water as working fluid. Fig. 2 shows an example of transient displacement of wire at the measurement point *M* and machining voltage. They are obtained just after the voltage is supplied to the gap between the wire electrode and the workpiece. In Fig. 2, *T*₁ shows the time when voltage is supplied and *T*₂ shows the time when an electrical discharge occurs. In the wave form of displacement, it can be observed that part A and B are affected by supplied voltage and discharge, respectively. There is a certain delay in the displacement wave form from the change of voltage. This is due to the wave propagation from the discharge point to the measurement point. Judging from the direction of wire displacement, it can be considered that the wire is attracted toward the workpiece after *T*₁ and repulsed away from workpiece after *T*₂. Their forces which act on the wire just after the voltage is supplied will be discussed in the following sections.

3. Influence of electrostatic force on wire electrode

3.1 Transient response to supplied voltage

In the above experiment, a discharge occurs just after the wire starts moving towards the workpiece. In order to investigate only the influence of the supplied voltage on displacement of the wire, the gap distance *d_g* between the wire electrode and the workpiece is set so that it would not lead to discharge. Fig. 3 shows displacement of the wire and amplitude of voltage just after stepwise voltage is supplied. In Fig. 3, the response of wire can be observed as an indicial response to the supplied voltage. The displacement of the wire cannot be observed in case of the experiment with the gap of air. As the dielectric constant of air is one-eightieth as that of water, this force which drives the wire can be considered to be electrostatic

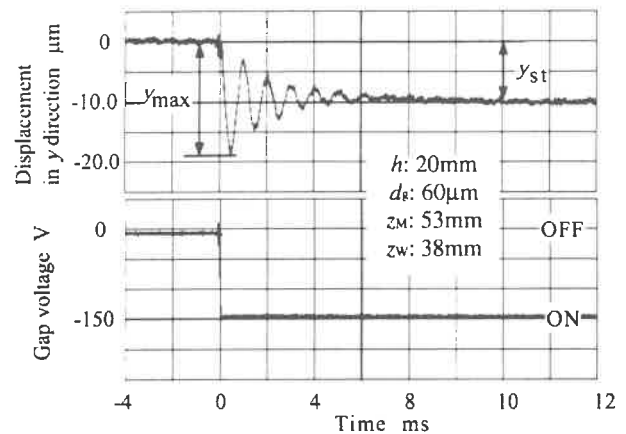


Fig. 3 Transient response of wire to stepwise voltage

force.

3.2 Discussion on electrostatic force

In order to evaluate the electrostatic force quantitatively, the final displacement in steady state *y_{st}* shown in Fig. 3 will be compared with the displacement calculated on a model of electrostatic field. In the calculation of the electrostatic force, the wire and the workpiece are assumed to be a cylinder and a plane, respectively. The influence of edge portions at the top face and the bottom face of workpiece can be neglected. Then, a static displacement *y_{st}* is calculated as a deflection of the wire caused by the electrostatic force which acts uniformly on the wire in the region limited to the thickness of the workpiece.

The relation between calculated wire displacement and the gap distance *d_g* is shown in Fig. 4 with the measurement results for different supplied voltages and for different thicknesses of workpiece. The observing height *z_m* for workpiece thickness *h*=1mm is 48mm and 53mm for *h*=20mm. The height of the workpiece *z_w* is 38mm for both cases. In Fig. 4, the plotted points and the curves show the measured results and the calculated results, respectively. The measured results are fairly close to the calculated results. Therefore, the input to the wire system can be considered to be an electrostatic force just after the gap voltage is supplied. In this experiment, the electrostatic force acting between the wire and the plain was obtained by the observation of transient response of the wire displacement. Also the magnitude of electrostatic force during the machining of a groove can be obtained.

By looking at the transient part of displacement in fig. 3, from time 0 to the wave diminished to *y_{st}*, two following points are known. One is that the period of the vibration applies closely to the first normal mode. The other is that the maximum transient displacement *y_{max}* is approximately twice the value of final displacement *y_{st}*. Therefore, the displacement can be considered to show

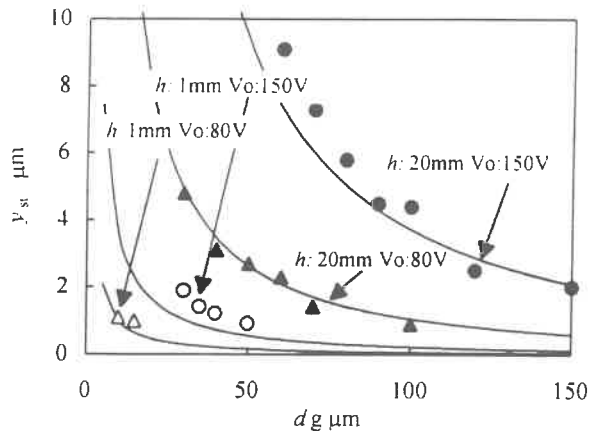


Fig. 4 Static displacement of wire caused by electrostatic force

the transient response to the step function of supplied voltage. Under the condition affected by an electrostatic force, the instantaneous decrease of the gap distance, right after voltage is supplied, becomes a primary factor which causes a discharge.

4. Vibration of beam under constant tensile force

In this chapter, the force which acts on the wire caused by a discharge is discussed. Just after the discharge occurs, the material of the workpiece melts at the discharge point and fluid surrounding the discharge point vaporizes. In this moment, the pressure caused by the vaporization removes the melting portion and then a discharge crater generates. Since the acting period of the pressure is shorter than the period of lower vibration mode, the pressure can be considered as an impulsive force. By looking at the period of part B in Fig. 2, it can be seen that the response wave to the impulsive force after discharging is different from the response wave form assumed in a vibration system of a string. The bending stiffness on the wire vibration is taken into consideration for the investigation of vibration system and the impulsive force.

4.1 Equation of wire vibration

It is supposed that the wire vibration system during machining is to be comprised of an elastic beam under a certain tensile force, damping force and an external force $F(z, t)$. The equation of motion in this vibration system can be written by

$$EI \frac{\partial^4 y}{\partial z^4} - T \frac{\partial^2 y}{\partial z^2} + \rho A \frac{\partial^2 y}{\partial t^2} + c \frac{\partial y}{\partial t} = F(z, t) \quad (1)$$

where $y(z, t)$ is the displacement of the wire, E is Young's modulus, I is the moment of inertia, ρ is the density, A is the cross-section area and c is the damping

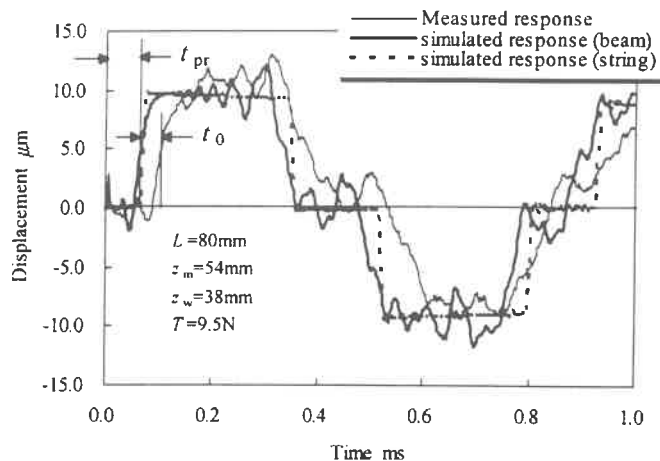


Fig. 5 Simulated response and measured response to impulsive force

coefficient under the consideration of water viscosity. The first term of equation (1) shows the force affected by stiffness. If the stiffness can be disregarded, it will be an equation for a string. From now on, the vibration system in which stiffness of the wire is considered is called a "beam model" and the vibration system in which stiffness is disregarded is called a "string model".

4.2 Numerical simulation of the wire

The numerical simulation based on equation (1) is calculated using the finite differences method under the condition that both ends are fixed.

It is assumed that influence of wire winding rate (0.145m/s) on the displacement response can be neglected because the wire winding rate is much slower than the propagation speed (187m/s).

Regarding the above vibration system, it is assumed that the impulsive force caused by a single discharge acts on a wire very locally. It is defined that a rectangular impulse of a magnitude F_0 and period t_i acts on an element of the wire at the height of the workpiece ($z_w=38\text{mm}$). The response of the wire displacement to the impulse is observed at the observing height ($z_m=54\text{mm}$). The initial conditions are given $y=0$ for all elements. In the numerical simulation, a distance interval of $\Delta z=0.1\text{mm}$ in z -direction and a time interval of $\Delta t=10\text{ms}$ were used.

4.3 Estimation of impulse caused by a discharge

In order to evaluate the availability of the model for the wire vibration, first, the simulated results are compared to the measured wave forms of the response to a single discharge. Then, an impulse caused by a discharge will be determined so that the calculated response corresponds to the measured response. The conditions of the simulation are shown in Table 2. In the experiment, the workpiece ($h=1\text{mm}$) thickness and the magnitude of supplied voltage

Table 2 Parameters in simulation of wire vibration

Young's modulus: E	N/m^2	9.6×10^{10}
Moment of inertia: I	m^4	7.854×10^{-17}
Tension of wire: T	N	9.5
Density: ρ	kg/m^3	8.53×10^3
Cross-section area: A	m^2	3.14×10^{-8}
Damping coefficient: c	m/(kg s)	0.05
Distance between two support guides: L	m	0.08

(80V) is set in order to disregard the influence of an electrostatic force.

The response in the experiment and the simulation are shown in Fig. 5. In the figure, the thin solid line, the thick solid line and the dotted line show the experimented result, the simulated result of the beam model and the simulated result of the string model, respectively. The simulated results of the beam model and the string model show that the wave form of higher frequencies and the reflected waves are represented clearly in the beam model. Thus, the condition of the supporting guides and the impulse which acts on the wire can be evaluated using the method above.

Since the observing point is offset from the discharge point, the simulated response at the observing point will not exactly be synchronized to the instance of the discharge meaning there is a time in delay t_{pr} as shown in Fig. 5 for the wave propagation. In the experiment, the observed response delays furthermore $-t_0$ as shown in Fig. 5. This is due to the time required to generate pressure on the wire after the discharge.

5. Consideration of transient behavior of wire electrode

As a result, after the gap voltage is supplied, the transition of wire behavior (a) to (d) in Fig. 6 follows the below:

- (1) At the moment the voltage is supplied stepwise between a wire electrode and a workpiece, the wire is attracted to the workpiece. The transient response of the wire represents the indicial response to electrostatic force which is distributed uniformly on the wire.
- (2) As the result of the attracting force, the gap between the wire and the workpiece becomes narrow. It leads to the occurrence of discharge.
- (3) Just at the occurrence of discharge, the attractive force disappears. After a certain interval, an impulsive force acts on the wire very locally, and the wire is deformed locally where the impulsive force is received.
- (4) The wave of displacement propagates toward the both ends with time. Thus, the displacement of the wire is observed as an impulse response, with the initial

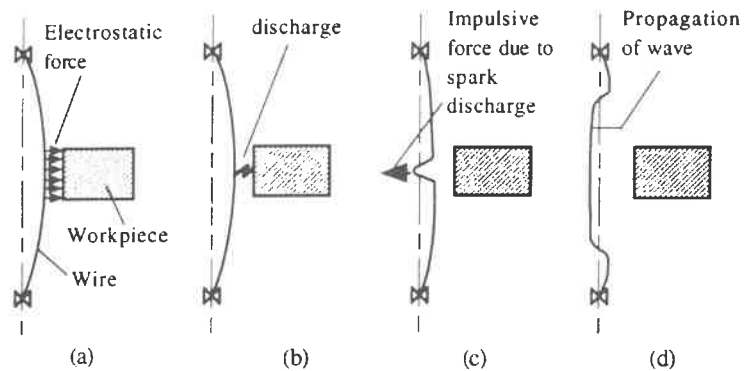


Fig. 6 Typical behavior of wire electrode in single discharge

condition of the wire just before the discharge occur.

6. Conclusions

In this paper, the influence of pressure caused by each discharge and the influence of an electrostatic force on wire displacement are investigated under a mechanical model supposing the tool electrode to be a beam under a certain tensile force. The obtained conclusions are as follows:

- (1) The response in experiments is fairly close to the impulse response simulated with the beam model of the wire.
- (2) Owing to small gap between electrode and workpiece, this electrostatic force acts on the wire electrode as a uniformly distributed load along the wire. The displacement of the wire induced by the electrostatic force is considered as the indicial response to applied stepwise voltage. The initial position for the impulse response to the following discharge is given by the displacement.

These considerations of the wire response to the exciting force caused by a single discharge helps us to investigate the primary factors of the wire vibration by continuous discharges in wire-EDM. Consequently it will lead to the precision machining in the future.

Acknowledgment

The authors wish to thank to Mr. Tamio Takawashi of Mitsubishi Electric Corp. for his helpful suggestions.

References

- [1] K.Horio, et al.: The study of wire behavior in wire EDM (1st report) -Wire vibration and discharge phenomena-, J. Jpn. Soc. Electrical Machining Eng., VOL.14, No.28, pp.30-41(1981)
- [2] W.Dekeyser: Knowledge-based system for wire-EDM, Ph.D. Thesis, K.U. Leuven, Belgium, pp.6-25 (1988) 6
- [3] H.Yamada, et al.: Modal analysis of wire electrode vibration in wire-EDM, Int. J. Electrical Machining, No. 2, pp.19-24(1997)

THREE-DIMENSIONAL MICRO FORMING USING METAL JET

Katsumi Yamaguchi, Katsuhiko Sakai, Takashi Yamanaka, Takao Hirayama

Department of Mechanical Engineering, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, 464-01 JAPAN

1. Introduction

In recent years, various micromachining processes have been developed to meet the demands for micro machines. It is well known that small sensors or micro parts are produced by a silicon fabrication process [1][2]. LIGA, laser assisted etching and 3-dimensional photo fabrication processes have also been developed to produce various microcomponents [3]–[7]. However, 3-dimensional parts cannot be formed by these processes, and the material is limited to silicon or plastics. In this study, the authors developed a method to form micro parts by a metal jet. This method is similar to an ink jet printer. A thin layer is first formed by molten metal jet ejected through a nozzle and a second layer is formed on the first layer. The repetition of this process produces the 3-dimensional parts. In this research, the basic performance of the metal jet head is examined, and some examples of 3-dimensional forming and a 3-dimensional circuit made by fusible alloy and Pb-Sn alloy are shown.

2. Principle of ejecting metal-jet drops

Figure 1 is an illustration of the metal-jet head used. A metal drop is ejected from the nozzle by a piezoelectric actuator. This device is very different from an ink jet printer head in various points. The density of the metal, for example, is much greater than that of ink, and the metal surface tension and metal-to-nozzle wettability are different from those of ink. Further, the melting point of alloy metal is too high for the layer type piezoelectric actuator. Thus, great care was taken to avoid these problems in the design of the metal jet head. For example, the actuator is protected against thermal damage by insulators and heat radiator.

The metal drops was generated as follows. The metal alloy is melted in an electrically

heated tank which is mounted above the nozzle, and the molten alloy is supplied to the chamber by gravity. The chamber has a diaphragm which connects to the piezoelectric actuator. This actuator is driven by the signal from a pulse generator through a high-capacity amplifier. The nozzle diameter is 0.2 mm and the metal drop ejected is about 0.30 to 0.40 mm in diameter. In this experiment, fusible alloy (Bi-Pb-Sn-Cd-In) and solder (Pb-Sn alloy) were used with a melting point of 47°C and 183°C, respectively.

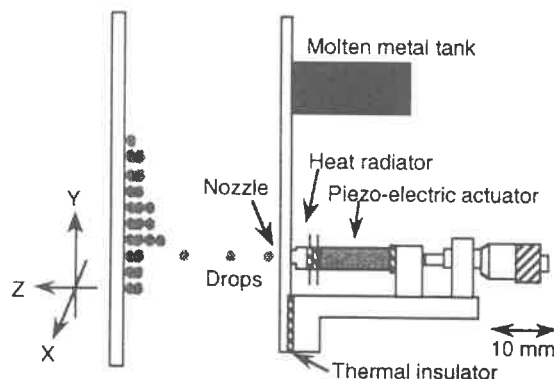


Fig. 1 Schematic illustration of metal-jet head and principle of 3-dimensional forming

3. Performance of metal-jet head

Voltage, pulse frequency applied to the actuator and the temperature of molten metal are the parameters greatly affecting the size, speed, shape, stability and reproducibility of ejected drop or dot on the base plate. Therefore, suitable conditions have to be found to produce a model for good accuracy. First, the shape of a flying fusible alloy drop is investigated. Figure 2 shows a continuous shot of the drop observed by a high-speed video system. The time interval of shots was 0.001 second. The drop seems to be oval and stable during the flight. One drop is ejected by one pulse signal. From these frames, the initial speed of drops was estimated, and shown in Fig. 3. The speed of drops can be controlled by the voltage applied to the actuator. Second, the accuracy of "printing" on the base plane was studied. Metal dots are ejected in line onto the base plane. Figure 4 is an image of the adhered dots under scanning electron microscope. The size and intervals of the dots are almost uniform. For example, the mean diameter of drops is $403 \pm 14 \mu\text{m}$, and positional error is within $\pm 13 \mu\text{m}$.

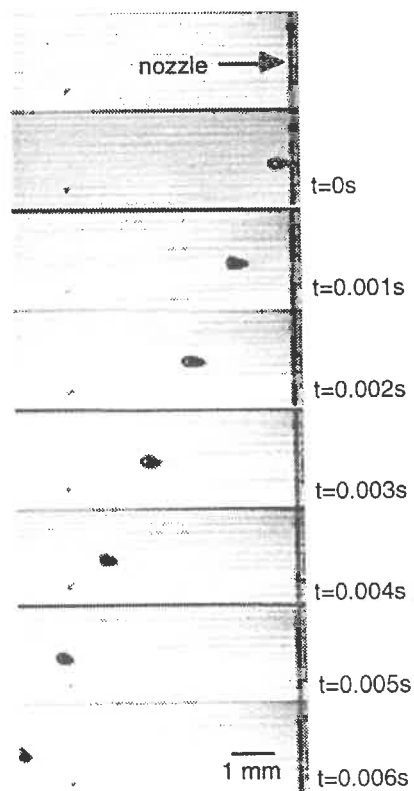


Fig. 2 Flight of a metal drop taken by high speed video system.

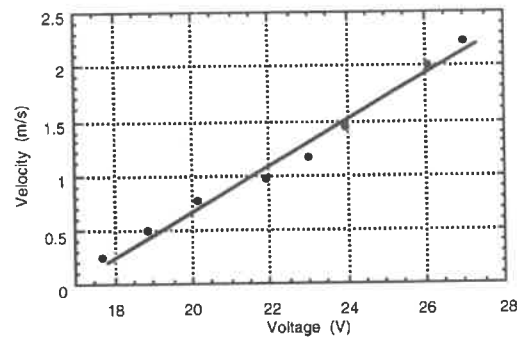


Fig. 3 velocity of metal drop

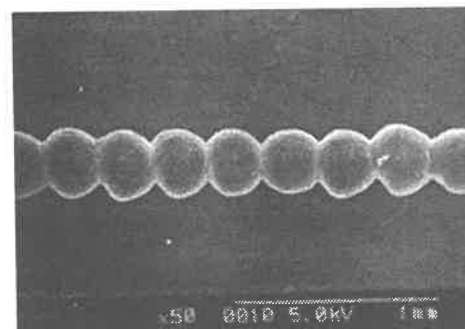


Fig. 4 SEM image of the metal dots on the base plane

4. 3-dimensional forming

Typical examples of a 3-dimensional model made by fusible alloy and solder are shown in Fig. 5 and 6. Models are formed by the layer-by-layer technique described in the previous section. Figure 7(a) presents an SEM image of the model surface, which is seen to have many small dimples. Although a few vacancies remain in the cross section shown in Fig. 7(b), the metal dots are very densely packed. To confirm this, the bulk density of the models was measured by the Archimedes method. The bulk density changes from 8.4 to 9.2 g·cm⁻³ as shown in Fig. 8. The density of the fusible alloy is 9.4 g·cm⁻³, so the maximum packing rate reaches 98%.

Figure 9 shows the coil of solder formed in the transparent resin. The alloy dots are connected with each other to assure electric conductivity. Therefore, this model can be used as an electrical circuit. This method can be applied to the production of the functional gradient material. A typical sample is shown in Fig. 10. The concentration of the metal and resin changes continuously in the middle of the model.

It can be easily recognized that this process has a wide variety of applications in the forming of micro components, electric circuits and functional gradient material.

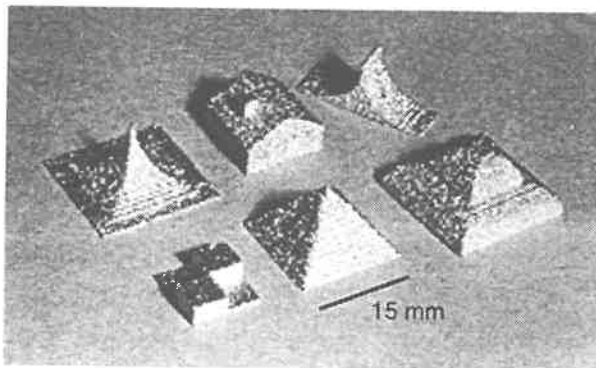


Fig. 5 Typical examples of 3-dimensional model made by fusible alloy

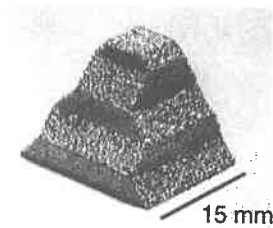


Fig. 6 3-dimensional model made by solder

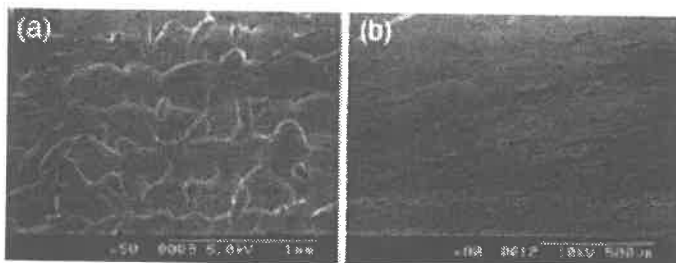


Fig. 7 (a) SEM image of the layer surface and (b) cross-sectional image of the specimen

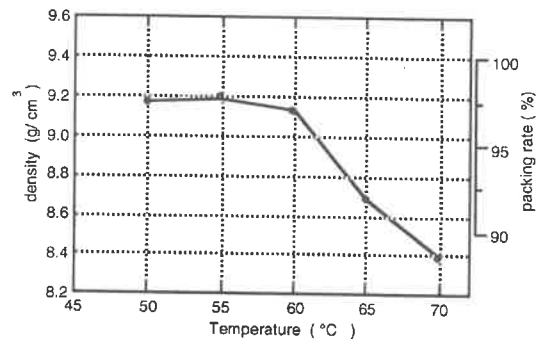


Fig. 8 Bulk density of the models

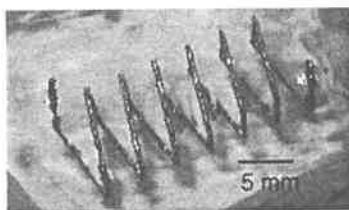


Fig. 9 Coil in the transparent resin

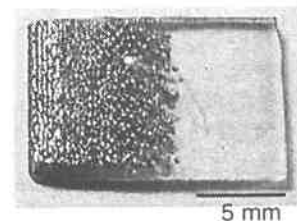


Fig. 10 Functional gradient material

5. Conclusion

1) A metal-jet prototype system was developed and its performance was shown. 2) 3-dimensional models, electric circuits and functional gradient materials were successfully produced. 3) The highest packing rate of the metal is 98%.

6. Acknowledgments

The authors would like to thank Mr. K. Nunome for setting up the experimental equipment and Mr. M. Yamasaki for cooperation in conducting the experiments.

7. References

- [1] M. Mehregany, K. J. Gabriel and W. S. N. Trimmer, Integrated Fabrication of Polysilicon Mechanisms, *IEEE Trans. Electron Devices*, 1988, **35**(6), 719.
- [2] L. S. Fan, Y. C. Tai and R.S. Muller, Integrated Movable Micromechanical Structures for Sensors and Actuators, *IEEE Trans. Electron Devices*, 1988, **35**(6), 724.
- [3] W. Ehrfeld, The LIGA Processes for Microsystems, *Micro System Technologies 90, Proceedings of 1st International Conference on Micro Electro, Opto, Mechanic System and Components*, 1990, 521.
- [4] J Mohr, C. Burbaum, P. Bley, W. Mentz and U. Wallrabe, Movable Microstructures Manufactured by the LIGA Processes as Basic Elements for Microsystems, *Micro System Technologies 90, Proceedings of 1st International Conference on Micro Electro, Opto, Mechanic System and Components*, 1990, 529.
- [5] H. Guckel, K. J. Strobis, T. R. Christenson, J. Klein, S. Han, B. Choi, E. G. Lovell and T. W. Chapman, Fabrication and Testing of the Planer Magnetic Micromotor, *Journal of Micromechanics and Microeng.*, 1991, **1**(3), 135.
- [6] B. Frazier, J. W. Babb, M. G. Allen and D. G. Taylor, Design and Fabrication of Electroplated Micromotor Structures, *Proc. of ASME Winter Annual Meeting.*, 1991, **DSC32**, 135.
- [7] K. Yamaguchi, T. Nakamoto and P. Abraha, Consideration on the Optimum Condition to Produce Micromechanical Parts by Photo Polymerization using Direct Focused Beam Writing, *Proc. of 6th International Symposium on MHS*, 1995, 71.

FABRICATION AND THERMAL ANALYSIS OF A BELLOWS TYPE GAS-LIQUID PHASE-CHANGE MICRO-ACTUATOR

Shigeo KATO and Tadashi SHINTANI

Nippon Institute of Technology,
Miyashiro, Saitama 345, Japan

1. Introduction

New inching mechanisms to inspect the human body or to repair thin pipings have been searched.

There are flexible inching mechanisms which are able to inch in pipings by imitating the creeping motion of earthworms. We have fabricated a metal bellows type gas-liquid phase-change micro-actuator for the engine of the flexible inching mechanism. When a liquid is vaporized, the volume of the phase-changed gas is more than 100 times that of the liquid. Gas-liquid phase-change actuators are thermal engines that utilize the large change in volume that occurs during a phase-change. [1],[2]

By using a metal bellows, a thin and long actuator has been obtained. Furthermore, the stretching and shrinking motion of the metal bellows actuator can be utilized for the inching motion of the flexible inching mechanism.

This paper presents a fabrication and thermal analysis of a metal bellows type gas-liquid phase-change micro-actuator.

2. Characteristics of the fabricated actuator

A cross section of the fabricated metal bellows type micro-actuator is shown in Fig. 1. The bellows is made of nickel. Its dimension is 6.8 mm in outer diameter, 3.8 mm in inner diameter, and 20.4 mm long. It has a volume of 536 μl for operating liquid. The operating fluid is non-chlorine perfluorocarbon ($\text{C}_5\text{H}_{11}\text{NO}$), because its vaporizing temperature, though higher than the human body, is still only 50°C and its latent heat for vaporization is only 122 that of water. Consequently, the energy needed for vaporization is very low. The operating liquid is heated by a heater 0.1-mm thick made of coiled constantan wire with an electrical resistance of 12 Ω . The elasticity of the bellows is almost linear in the displacement less than 2.5 mm. The gauge pressure in the bellows is 0.04

MPa when the displacement of the bellows is 2.5 mm.

An electrical power of 5.88 W is supplied

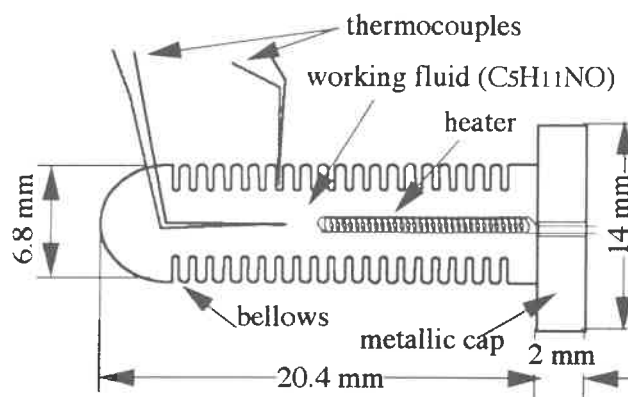


Fig. 1 Metal bellows type gas-liquid phase-change actuator

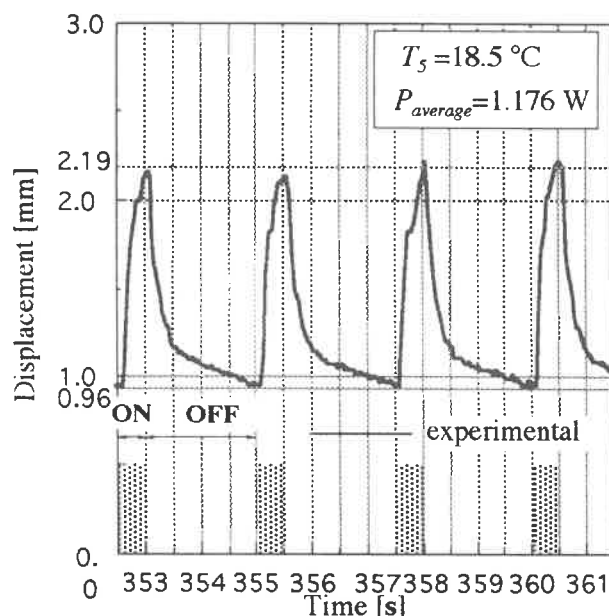


Fig. 2 Continuous stroke motion
(Average power is 1.176W)

within 0.5 seconds to the heater and it is cut in the interval of 2 seconds. This operation is repeated. The maximum displacement of the piston gradually increases. After 60 seconds, the maximum displacement reaches 2.19 mm and the minimum displacement reaches 0.96 mm. This motion is shown in Fig. 2.

Thus, this actuator was confirmed to have a reciprocating motion with a 1.23-mm stroke at an average power of 1.176 W.

Next, the electric power was given continuously. The displacement of the bellows increases according to the power. However, if the power decreases to 1.17 W, the stroke of the piston freezes at the length of 0.81 mm. The performance of the actuator is considered to be thermally critical. In this case, the ambient temperature (T_5) is 18 °C, the temperature of the outer wall of the bellows is 29.2 °C, and the temperature of the operating liquid at 1 mm from the heater is 33 °C.

3. Mathematical model of the actuator

In order to simulate the actuator, we needed to create a mathematical thermal model of the actuator. A model of the actuator is shown in Fig. 3. The temperature distribution is very complicated. However, we simply considered four thermal elements receiving the heat and their temperature changes. These are (1) the heater, (2) the operating liquid near the heater (boiling temperature zone), (3) the operating liquid near the casing, (4) the bellows and a metallic cap. If we select only one liquid for the operating fluid, we cannot explain why vaporization starts 0.2 seconds after power is supplied. The heat capacity, the temperature, and the heat transfer area of the element- i are C_i (J/K), T_i (°C), and A_i (m²), respectively. The heat flow rate and the heat transfer coefficient from element- i to element- $(i+1)$ are q_i (W), and h_i (W/m² K), respectively. The heat flow rate to the heater is q_H (=5.88 W). The heat flow rate for vaporizing is q_v (W). The heat flow rate when the vaporized gas is re-liquefied and the liquefying area are q_L (W) and A_L (m²), respectively. Simultaneous differential equations of the thermal model are solved by the Runge-Kutta method.

3.1 Under power supply (from start to vaporization)

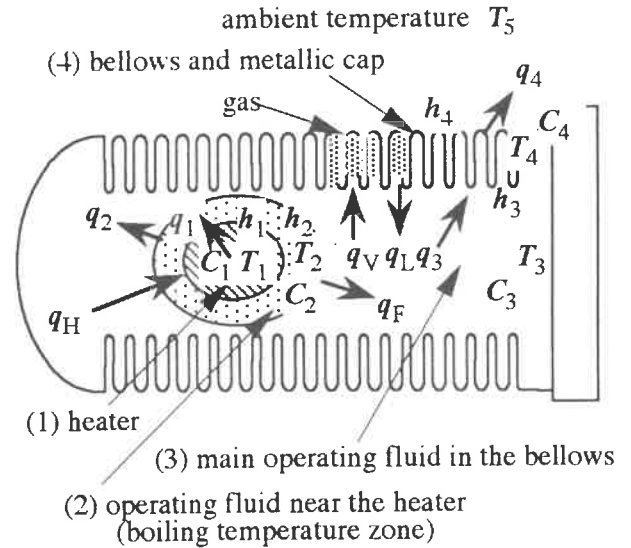


Fig. 3 Thermal model of the metal bellows type gas-liquid phase-change actuator

When power is supplied to the heater, the operating liquid near the heater quickly vaporizes. For the period from the start till vaporization occurs, the heat balances of the four thermal elements are represented by next four differential equations. t represents the operating time.

$$C_1 \frac{dT_1}{dt} = q_H - q_1 = q_H - A_1 h_1 (T_1 - T_2) \quad (1)$$

$$C_2 \frac{dT_2}{dt} = q_1 - q_2 = A_1 h_1 (T_1 - T_2) - A_2 h_2 (T_2 - T_3) \quad (2)$$

$$C_3 \frac{dT_3}{dt} = q_2 - q_3 = A_2 h_2 (T_2 - T_3) - A_3 h_3 (T_3 - T_4) \quad (3)$$

$$C_4 \frac{dT_4}{dt} = q_3 - q_4 = A_3 h_3 (T_3 - T_4) - A_4 h_4 (T_4 - T_5) \quad (4)$$

3.2 Under power supply (after the vaporization)

When the temperature of the operating liquid near the heater reaches the boiling temperature ($T_b=50$ °C), vaporization is assumed to start from the sidewall of the heater. Eqs. (2) and (3) are changed to Eqs. (5) and (6), respectively.

$$C_2 \frac{dT_2}{dt} = 0 \quad (5)$$

$$C_3 \frac{dT_3}{dt} = q_2 - q_3 + q_L$$

$$= (A_2 h_2 + A_L h_L)(T_b - T_3) - A_3 h_3 (T_3 - T_4) \quad (6)$$

$$q_1 - q_2 = A_1 h_1 (T_1 - T_b) - A_2 h_2 (T_b - T_3) = q_v + q_F \quad (7)$$

The zero in Eq. (5) shows that temperature T_2 maintains the boiling temperature, and $(q_1 - q_2)$ is obviously not equal to zero as Eq. (7). The term q_F in Eq. (7) is the heat flow rate that is used to spread the vaporizing temperature zone (the operating liquid near the heater). Supposing that β is a constant less than 1, it is assumed that terms q_v and q_F can be represented as follows.

$$q_v = \beta (q_1 - q_2) \quad (8)$$

$$q_F = (1 - \beta)(q_1 - q_2) \quad (9)$$

Increasing of the vaporizing temperature zone is represented as

$$(T_b - T_3) \frac{dC_2}{dt} = q_F = (1 - \beta) (q_1 - q_2) \quad (10)$$

The mass per unit of time of the vaporizing fluid (kg/s) is represented as

$$m_v = \frac{q_v - q_L}{H_b} \quad (11)$$

where H_b is the latent heat for vaporization (kJ/kg K) of the operating fluid.

4. Estimation of the heat transfer coefficients

Using the heat transfer areas shown in Table 1 and critical power of 1.17W, some heat transfer coefficients have been estimated.

When power of the heater is critical, the heat transfers between each thermal element are static. Consequently, following equations are obtained.

$$q_1 = q_2 = q_3 = q_4 = q_H = 1.17 \text{ (W)} \quad (12)$$

$$A_1 h_1 (T_1 - T_2) = q_H \quad (13)$$

$$A_2 h_2 (T_2 - T_3) = q_H \quad (14)$$

Table 1 Simulation data

Heat Transfer Area	$A_1=63.5 \times 10^{-6} \text{ m}^2$ $A_2=78.4 -$ $A_3=1377$ $A_4=2025$ $A_{L2}=270$
Heat Capacity	$C_1=5.97 \times 10^3 \text{ J/K}$ $C_2=1.27 \times 10^{-3} -$ $C_3=-0.951$ $C_4=2.546$
Heat Transfer Coefficient	$h_1=648 \text{ W/m}^2 \text{ K}$ $h_2=648$ $h_3=224$ $h_4=52$ $h_L=25$
Ratio of Heat Flow Rate for Vaporization	$\beta=0.06$

$$A_3 h_3 (T_3 - T_4) = q_H \quad (15)$$

$$A_4 h_4 (T_4 - T_5) = q_H \quad (16)$$

We have considered that T_2 is at a boiling temperature of 50 °C as mentioned and T_5 is 18 °C. Under these conditions, we obtained that h_2 is 648, h_3 is 224 and h_4 is 52 (W/m² K). However, we can not obtain only h_1 by these equations.

Calculation of the heat transfer coefficient h_1

We calculated the heat transfer coefficient h_1 as natural convection from a 0.1-mm horizontal hot column. The Nusselt number Nu is represented as

$$Nu = 1.03 B_1 Ra^{0.25} \quad (17)$$

where B_1 is a constant (=0.46) depending on the Prandtl number (=12.14). Ra is the Rayleigh number, and is 27.9 in this case. By substituting these values into Equation (17), we have $Nu=1.108$. Finally, we obtain $h_1=648$ (W/m² K) from the Nusselt number.

The heat transfer area for liquefying A_L is an area ($A_{L2}=270 \text{ mm}^2$) if the displacement is as large as $H=2 \text{ mm}$. However, if the displacement of the piston decreases, the area where the gas and liquid come in contact with each other may

decrease due to the capillary phenomenon. If the displacement is zero, then area A_L may be zero. Consequently, we assume that area A_L is related to displacement H and is represented as

$$A_L = A_{L2} \left(\frac{H}{H_2} \right)^N \quad (18)$$

where H_2 shows a displacement of 2-mm.

The heat transfer coefficient for liquefying h_L is assumed to be $h_L = 25 (\text{W/m}^2 \text{K})$ and the value of N is 5, because this value enables the minimum displacement of the piston to reach $H = 0.96 \text{ mm}$.

5. Simulation of the displacement of the piston

We simulated the displacement of the actuator using the values $\beta = 0.06$, because this value enables the displacement of the piston to reach $H = 2.19 \text{ mm}$ of the maximum displacement.

The relationship between volume V (μl) of the vaporized gas and displacement H (mm) is represented as

$$V = 22.05 H \quad (19)$$

The relationship between the pressure in the rubber membrane P (MPa) and the displacement H (mm) is represented as

$$P = 0.101 + 0.016 H \quad (20)$$

The pressure and the volume are represented as

$$PV = M_v R T_b \quad (21)$$

where M_v and R are the mass of the gas in the rubber membrane and the gas constant of the operating fluid, respectively.

Because the molecular weight of $\text{C}_5\text{H}_{11}\text{NO}$ is 310, from Eqs. (19), (20), and (21), the relationship between the displacement H and the mass of the gas M_v (kg) is represented as

$$H (0.101 + 0.016 H) = 4.072 \times 10^5 M_v \quad (22)$$

If we know the mass of the gas, we obtain the displacement of the piston by Eq. (22).

The simulation of the continuous reciprocating movement of the piston is shown in Fig. 4 by the gray line. It is confirmed that

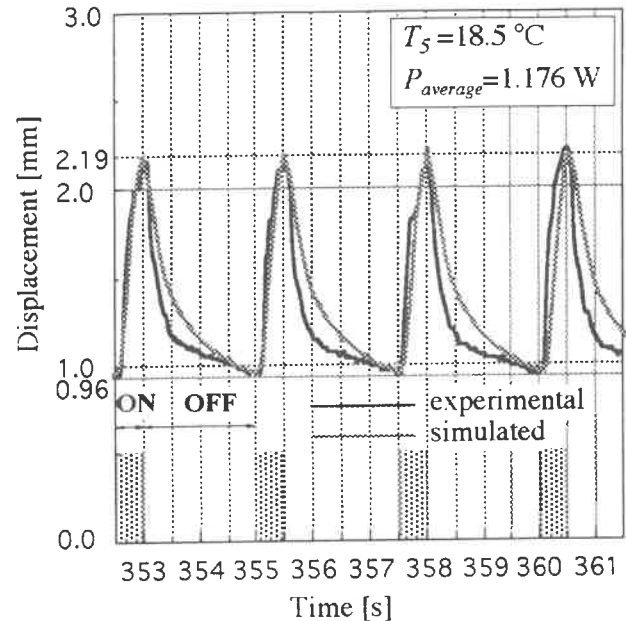


Fig. 4 Continuous stroke motion
(Average power is 1.176W)

continuous reciprocating movement of the piston with a 1.23-mm stroke can be simulated.

6. Conclusions

(1) We fabricated a metal bellows type gas-liquid phase-change micro-actuator that is 6.8 mm in diameter and 20.4 mm long, and that has a volume of 536 μl for the operating liquid. The operating fluid is $\text{C}_5\text{H}_{11}\text{NO}$.

(2) By dividing the operating fluid into two thermal elements, we could make a simple mathematical model for the gas-liquid phase-change micro-actuator.

(3) By experiment and simulation, the actuator was confirmed to reciprocate continuously with a 1.23-mm stroke at an average input power of 1.17W.

Reference

- [1] H. Mizoguchi, M. Ando, T. Mizuno, T. Takagi, and N. Nakajima: Proc. of Micro Electro Mechanical Systems '92, Design and Fabrication of Light Driven Micropump, pp. 31-36.
- [2] S. Kato: Proc of ASPE, Fabrication and Thermal Analysis of a Gas-liquid Phase-change Micro-actuator, pp. 332-335.

MINIATURIZED POSITIONING DEVICE FOR PRECISE ROTARY MOTION BY USE OF PIEZOELECTRIC ACTUATORS

Peggy Dörner

Fraunhofer Institute for Applied Optics and Precision Engineering
Jena, Germany

Introduction

With increasing miniaturisation, more accurate positioning systems for precise handling and for examining components for quality are required. Such high accuracy requirements can be met by piezoelectric actuators, easily producing displacements on the order of submicrons by voltage control. However, in applying such actuators it is difficult to realize a rotary motion.

At the Fraunhofer Institute in Jena new positioning devices for precise rotary motions have been developed, driven by piezoelectric multilayer actuators. The deformation of one actuator generates a precise linear motion, which is transformed to rotary motion by an elastic hinge. This elastic hinge constitutes the pivot point of the whole positioning system. Due to the use of such a mechanism an angular resolution of 0.025 seconds of arc could be proved for the first developed device and for a second miniaturized device with a volume of 2.3 cm^3 , as well.

First design

For the first design, work was centered on the development of a new piezoelectric positioning system with high accuracy for generating a slow step by step rotary motion of a rotary table (rotor) /1/.

The general idea contains three driving elements: the first (1), which is not necessary in all cases, positions the rotor (5) at a certain height; the second (2), brings the stator (4) into contact with the rotor (5), and the third (3), generates the true rotary motion. All driving elements used are piezoelectric multilayer actuators. The first driving elements (1), for positioning, the rotor at a fixed height, exist as three rectangular columns protruding from the stator base plate to surround the adjustable holder of the stator (4). Each contains a piezoelectric actuator for varying the height of the column. The displacement of each piezoactuator is transmitted by an arrangement of elastic hinges, which enlarges the range for adjustment from $28 \text{ }\mu\text{m}$ to $200 \text{ }\mu\text{m}$. A similar mechanism is used for bringing contact between the adjustable holder of the stator (4) and the rotor (5). The maximum adjustment of the stator height is $200 \text{ }\mu\text{m}$, as well.

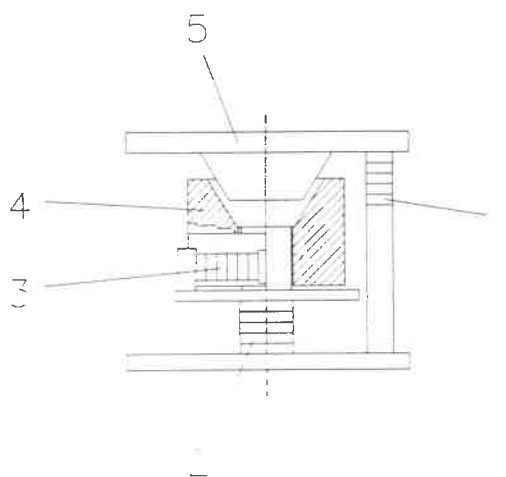


Fig. 1: Structure

The horizontally arranged piezoactuator (3) generates a variable linear motion, transformed by an elastic hinge into a rotary motion of the adjustable holder of the stator (4).

For driving this positioning system, a special electronic device has been developed, consisting of two parts: a logical part for generating the chronological sequence and an analogical part for driving each actuator.

The measured minimum rotary motion is 0.025 seconds of arc and the maximum angular step is 104 seconds of arc. For measuring the rotational motion, two plane reflectors are mounted on the rotor, the measuring reflector and the reference reflector. In addition, the measuring set-up contains a laser, an interferometer, and an analyser for calculating the angle around which the two reflectors are rotated. The angle is determined from the difference in path length of the light beams.

Design of the miniaturized positioning system and its improvements upon the first design

Because the first design has some disadvantages in industrial application: for example, large volume of the whole, relatively slow operation and restriction of operation to only one position, there is a demand for a miniaturized positioning system, which can work in different positions.

The miniaturized positioning system is based on a stator which is surrounded by a tubular rotor. With exception of the piezoelectric actuators, the stator of the positioning system, depicted in figure 2, consists of only one piece having three layers; the middle layer for rotary motion generation and two outer layers for clamping,. To produce the layers the stator is given two horizontal slits, guaranteeing a clear separation of rotary motion and clamping function. The clamping layers are identically constructed and include an extra structure, due to the manufacturing of the three layers as one mechanical piece, which however has no influence on clamping function. The whole stator shape, including the microstructures, is produced by electrodischarge machining. By this process very complicated shapes can be manufactured with minimal tolerances. The diameter of the presented positioning system is 12 mm, but, in the area of the clamp shoes, it is some micrometers greater. The height of the stator is 10 mm and the volume of this miniaturized instrument with rotor is 2.3 cm³.

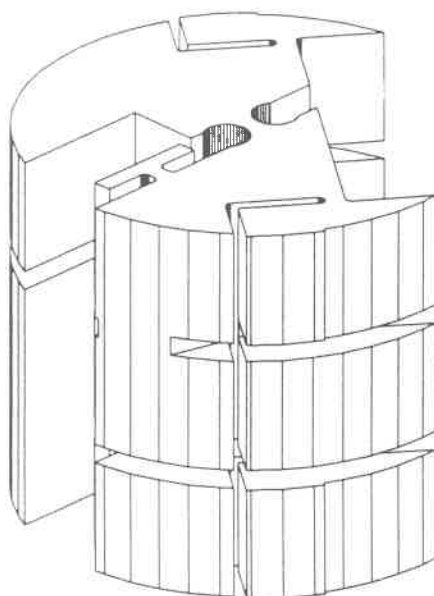


Fig. 2: Stator

Both clamps contact the rotor when there is no power consumption. Starting the operational sequence the piezoelectric actuator of the lower clamp is controlled by a voltage up to 150 V releasing this clamp from the rotor. Then an extension of the piezoelectric actuator in the middle layer causes a rotary motion of the upper layer which is still clamped to the rotor, rotating the rotor. To enable further rotation of the rotor, the lower clamp secures the position of the rotor again while the upper clamp releases. The piezoactuator in the middle layer can return back to its original position without rotating the rotor and the process can begin again. In this way a step-wise rotary motion of the rotor is generated without limiting range of rotation.

To avoid complicated micro-assembly techniques in general, the mechanical body of the positioning device consists of only one piece and there is no need for additional conventional bearings. The only assembly involves attaching the actuators. The simple structure of this miniaturized positioning device has the additional advantage of minimized sensitivity to vibrations.

Details of design

For generating the rotary motion a transfer of the linear motion of the piezoelectric actuator is necessary.

Elastic hinges give the possibility to change the type and direction of motion generated. In order to prevent tribological and assembly problems and to ease manufacturing with small tolerances, an elastic hinge has been chosen to constitute the pivot point of the whole positioning instrument. The measurements on the first instrument show, that the motion generation by means of piezoelectric actuators and the transformation and transmission of this motion by elastic hinges guarantees a very high angular resolution. Figure 3 presents the mechanism, which generates the rotary motion. The highest stiffness is along the vertical y-axis. The ratio between the width of the hinge to the diameter of the rounded end of the slit creating the hinge determines the ratio between the stiffness in the y-direction to the stiffness of rotation around the elastic hinge.

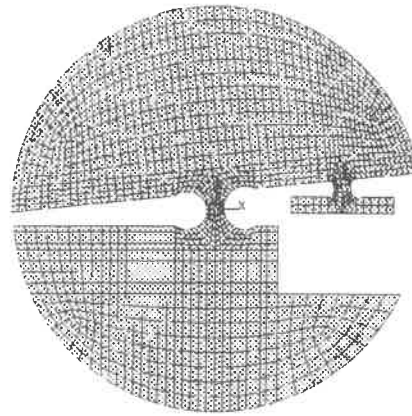


Fig. 3: Layer for generating rotary motion

The concept of clamping is based on the idea that at all times one clamp is active as in linear positioning systems like /2/ and /3/. In case of no power consumption both clamps clamp on to the rotor surrounding the stator. Therefore, the miniaturized positioning instrument is usable in any position.

Figure 4 shows the functional structure of one clamp containing a multilayered piezoelectric stack. When a voltage is applied the clamp shoes move inwards and are disengaged from the rotor. The pivot point for this motion is again an elastic hinge. The sort of clamping is influenced by the clamp shoes, the actuator and its distance to the pivot points and, as well, the surrounding rotor. The size of the gap between the clamp shoes and the rotor is very critical to the performance of the miniaturized positioning system.

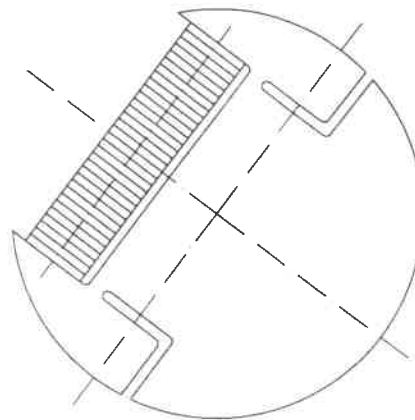


Fig. 4: Functional structure of clamp

Control

There are different possibilities for operational sequence. One will be explained here:

With no power consumption both clamps contact the rotor. After releasing clamp 1 the actuator for movement generation extends causing a rotary motion of clamp 2 and the rotor. Now clamp 1 contacts the rotor again and clamp 2 releases. In the next step, the contraction of the actuator for movement generation causes a rotary motion of clamp 2 without rotating the rotor. Such a cycle realizes a rotary step motion of the tubular rotor. The successive repetition of this cycle generates an unlimited step-wise motion of the rotor. The following figure shows the control for such a cycle.

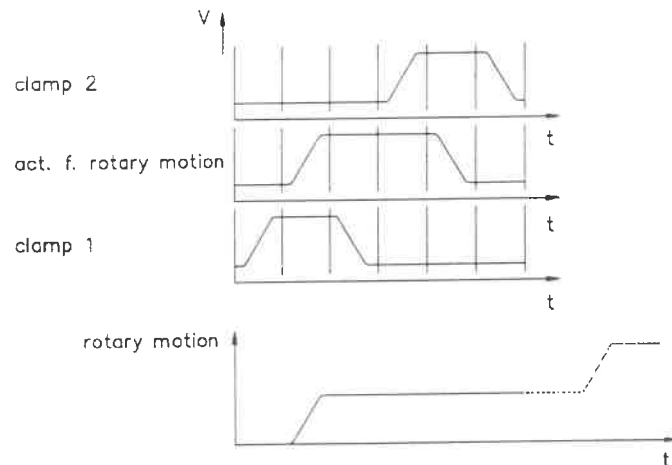


Fig. 5: Control of the actuators and generation of rotary motion

Measurement

The measuring set-up contains a laser, an interferometer and an analyser. For measuring the angular resolution the piezoelectric actuator for generating rotary motion was controlled by a triangular voltage. The measurement results are shown in figure 6 on the right. Every step is 0.025 seconds of arc high, which reflects exactly the angular resolution of used measurement set-up. That means there is no measurable reduction of displacement resolution of piezoactuator by the mechanism with elastic hinge.

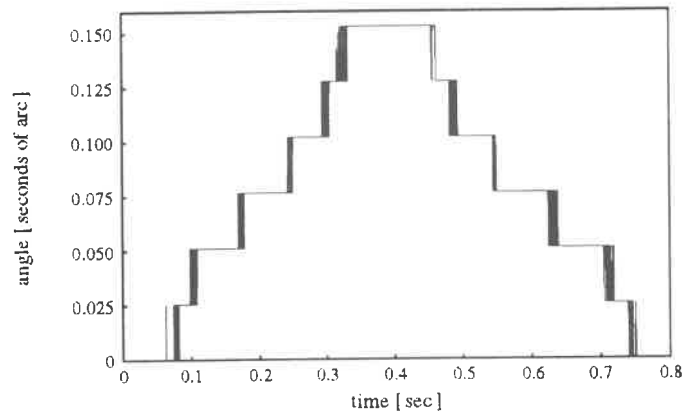


Fig. 6: Measurement of angular resolution

Conclusion

The miniaturized positioning device has been developed for high accuracy positioning of objects for handling and examination, for adjustment and quality control and for precise alignment of sensors. Due to the clamping without power consumption it is a long-time stable adjusting device. It can be coupled with linear positioning systems.

References

- /1/ Portius, P.; Wehrsdorfer, E.: Piezoelectrically Driven Rotation Stepper Motor for High Resolution Movings. Proc., 4th Int. Conference on New Actuators (Actuator 94), Bremen, June 1994, pp. 183 - 185
- /2/ Burleigh Instruments, US Patent 3,902,084
- /3/ Philips Electronics N.V., European Patent EP 0 592 030 A1

Dynamic Accuracy Monitoring for the Comparison and Optimization of Fast Axis Feed Drives

O. Zirn**, S. Weikert*

**Mikron SA Agno, Lugano, Switzerland

*Inst. of Machine Tools and Mfg.(IWF-ETH), Zürich, Switzerland

Abstract:

Based on the milling path verification via a 2-D incremental measuring grid, high speed machine tools as well as serially coupled high dynamic feed axis drives can easily be investigated and optimized. Supported by the Swiss machine tool industry, a method for the identification and the optimization of the accuracy and dynamic performance of fast axis feed drives has been developed. The exemplary comparison of a direct driven as well as a ball screw driven milling and engraving machine elucidates the obtainable improvements by this method.

Keywords:

Drive, Optimization, Measurement

1 Introduction

The successful application of the high speed machining process requires feed axis drives and machine structures capable of driving the tool at adequate feed rates along the milling path.

The efficient design and optimization of the machine tool must be based a model concerning all significant influences in the design phase and a suitable dynamic accuracy monitoring for the set-up and acceptance test.

Finite Element- (FE-) models [1] allow exact modeling but require suitable software tools that are not generally available in today's machine tool industry. The verification of the FE-models requires frequency analysis results from the prototype or comparable machine structures with known boundary conditions.

For the present investigation the application of the incremental 2-D-measuring grid in combination with a model of flexibly coupled discrete masses has proven to be advantageous. It yields useful results with moderate effort.

2 Modeling of high speed machine tools

Figure 1 shows the body structure of two exemplary state of the art high speed machine tools. Both machine types detach the X- and the Y-axis to obtain

optimum acceleration capabilities. This represents a basic boundary condition for the successful application of direct drives.

Serial machine principles for 3 or more machining dimensions require at least one cinematic chain of feed axes. In the presented examples (figure 1) the vertical Z-axis is attached on the slider (a)) or the column (b)) of the Y-axis. The Y- and Z-axis machine structure can't be realized with infinitive stiffness and thus must be considered as a system of flexibly coupled masses.

The division of the structure in two equal masses is a helpful approximation for a wide range of cases. The ball screw drive as well can be modeled as two flexibly coupled masses [4].

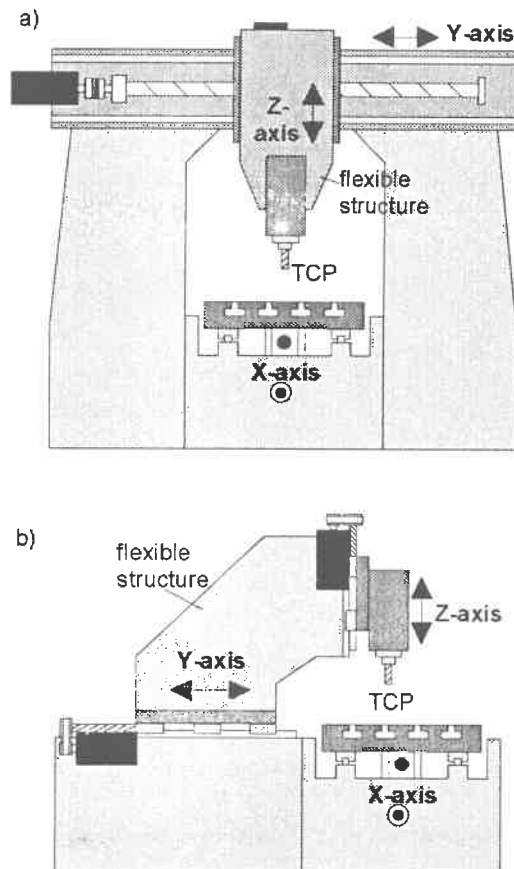


Figure 1: Typical axis arrangement of HSM machine tools with serial kinematics

A ball screw drive with a belt gearing can be modeled as a system of three flexibly coupled masses (see figure 2).

The complete model of the machine tool is shown in figure 2. The dynamic performance of the Z-axis will not be considered in this contribution.

The block diagram including the nonlinear influences of reluctance, friction and input saturation is shown in figure 3.

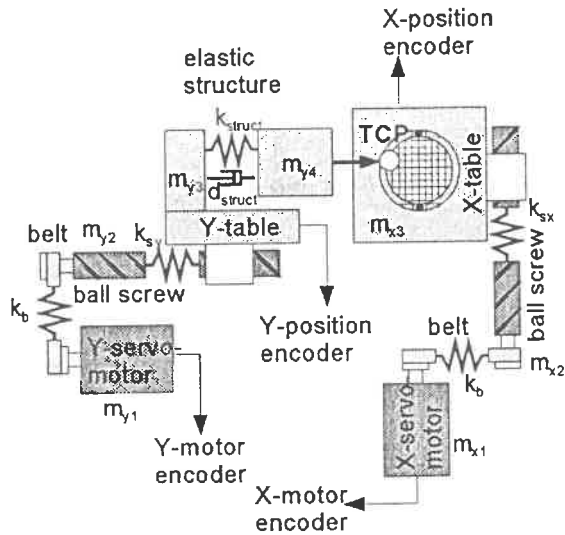


Figure 2: axis and flexible structure model

In this contribution an exemplary high speed milling and engraving machine tool, based on machine principle b) in figure 1 is investigated and optimized using the above (figure 2) shown machine model and the 2-D-measuring grid. The main technical data and the model parameters are shown in table 1. Most of the model parameters can be gathered from the components data sheets. The other parameters can be calculated or estimated basing on the machine's construction data.

parameter list	Y-axis	X-axis
table or structure mass m [kg]	250	50
stroke L [mm]	250	400
ball screw pitch h_s [mm]	10	10
ball screw stiffness k_s [N/m]	$154 \cdot 10^{15}$	$154 \cdot 10^{15}$
belt stiffness k_b [N/m]	$350 \cdot 10^{15}$	$350 \cdot 10^{15}$
structure eigenfrq. ω_{oS} [rad/s]	157	>1500
structure damp. d_{struct} [Ns/m]	9200	-
motor friction F_{R1} [N]	150	150
screw friction F_{R2} [N]	150	150
slider friction F_{R3} [N]	100	60
maximum force F_{max} [N]	8050	8050
reluctance force F_{rel} [N]	40	40
m_1 [kg]	319	319
m_2 [kg]	125	125
m_3 [kg]	138	150
m_4 [kg]	138	-
electrical time lag T_{elr} [s]	$2 \cdot 10^{-4}$	$2 \cdot 10^{-4}$
delay time T_{dead} [s]	$1.5 \cdot 10^{-4}$	$1.5 \cdot 10^{-4}$

Table 1: machine data and model parameters

Usually the relative damping of screw and belt is smaller than 5 % and therefore negligible. The machine structure damping can be higher and will be considered in the model.

The structure spring constant k_{struct} is much smaller than the spring constant of the belt k_b or the screw k_s . In this case the eigenfrequency ω_{oS} of the machine structure can be approximated by:

$$\omega_{oS} \approx \sqrt{k_{struct} \cdot \frac{m_{1y} + m_{2y} + m_{3y} + m_{4y}}{(m_{1y} + m_{2y} + m_{3y}) \cdot m_{4y}}} \quad (1)$$

where m_{1y} is the motor mass, m_{2y} is the mass of the half screw, m_{3y} is the half screw and the half structure mass as well as m_{4y} is the other half of the structure mass.

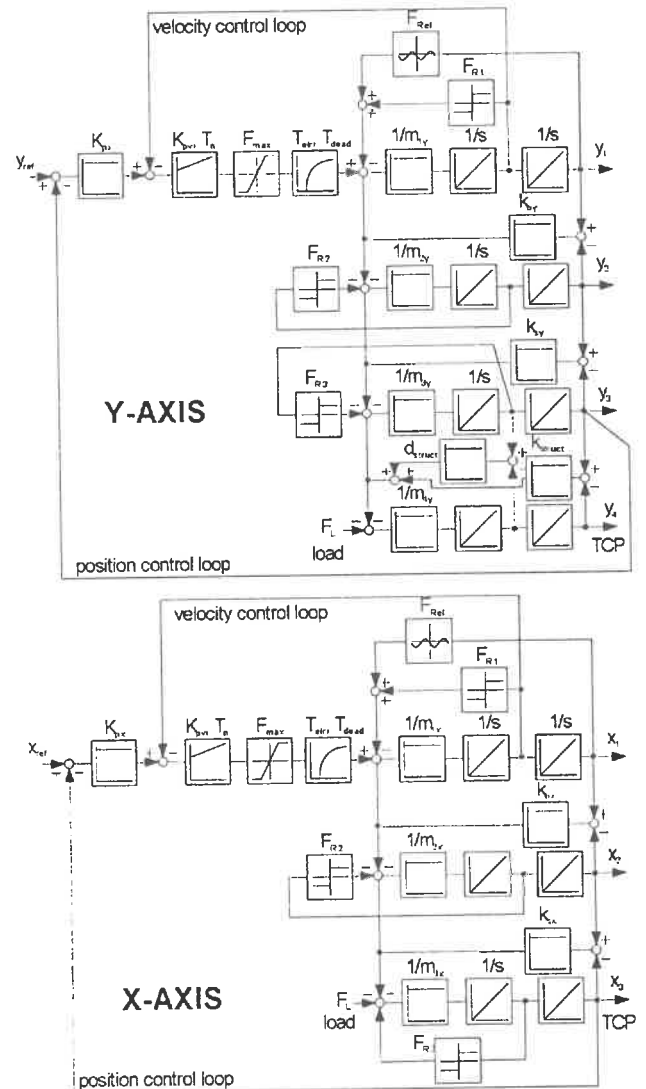


Figure 3: Block diagram of X- and Y-axis with ball screw drives

3 Analysis of tool trajectories

The measuring grid [2] (figure 4) is well suited for observing the relative planar motion of two bodies.

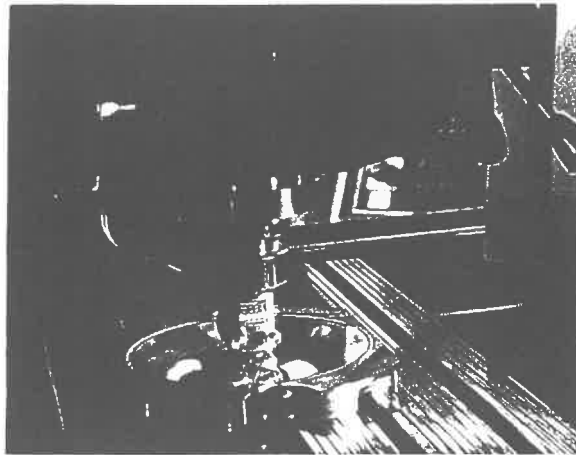


Figure 4: 2-D incremental measuring grid

The grid plate carries two line grids orthogonal to each other. A scanning head, mounted in the tool center point (TCP) also equipped with orthogonal line grids simultaneously lights the grid and observes the reflected light. The signal picked up is interpolated in 1024 intervals by the measuring circuit which results in a resolution of 5 nm. A PC is then used for storing and further processing the recorded data. The maximum speed is 24 m/min, which can be altered by reducing the resolution.

Figure 5 shows the measured tool trajectories at a YX-corner (upper figure) and a XY-corner (lower figure). The reference speed is 10 m/min with a reference acceleration limit of 1 m/s². The controller gains are the drive builders default values (table 2):

controller gains	Y-axis	X-axis
position gain K_{px} [1/s]	83	83
velocity gain K_{pv} [Ns/m]	350000	350000
integrator time constant T_n [s]	0.003	0.003

Table 2

This measurement shows clear that the controller gains of the machine tool axes are too high and lead to oscillations of the X-table and an overshoot of the Y-axis in combination with the flexible structure.

Computer based simulation with the nonlinear machine model in figure 3 allows a fast optimization of the controller gains (table 3).

controller gains	Y-axis	X-axis
position gain K_{px} [1/s]	50	50
velocity gain K_{pv} [Ns/m]	250000	160000
integrator time constant T_n [s]	0.003	0.003

Table 3

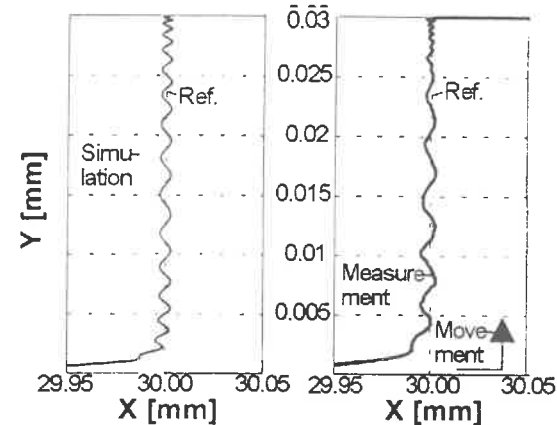
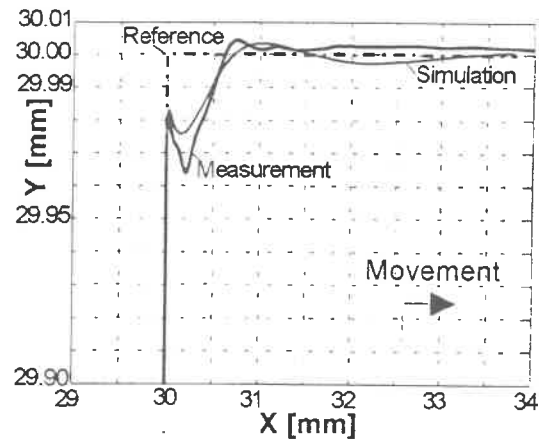


Figure 5: Comparison of the path deviation (measured and simulated tool trajectories)

Simulation results and the monitored path using the controller gains in table 3 are shown in figures 6a, 6b. The comparison of measurement and simulation shows good agreement. Deviations are caused by the different initial conditions of the simulation as well as estimation errors of parameters and the simplification. Although the nonlinear machine model (fig.3) is well suited for the design and the analysis of high speed machine tools in combination with the measuring grid.

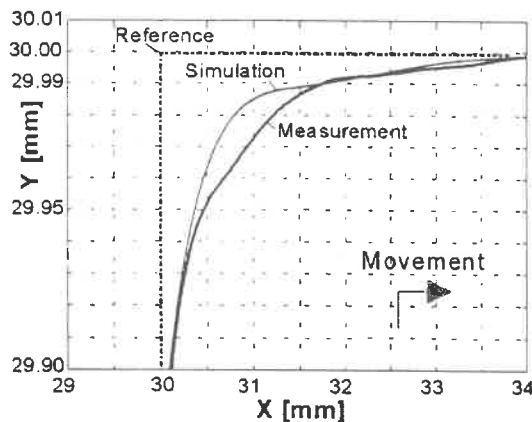


Figure 6a: TCP-path with moderate controller gains (2-D- grid measurement and simulation, 10 m/min reference speed)

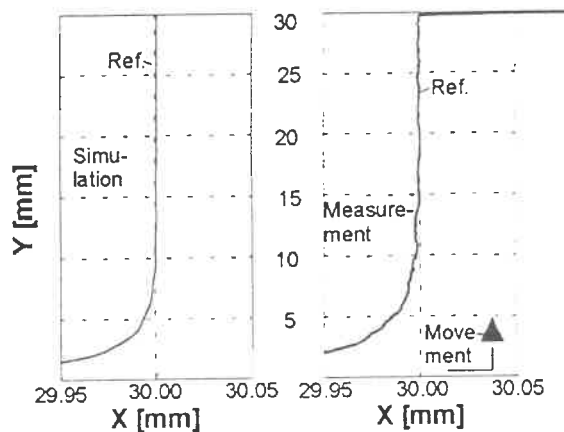


Figure 6b: TCP path with moderate controller gains (2-D-grid measurement and imulation, 10 m/min reference speed)

4 Direct Drive

The same milling and engraving machine type has been equipped with direct drives in the X- and Y-axis. The technical data and the model (figure 7) parameters are shown in table 4.

parameter list	Y-axis	X-axis
table mass m [kg]	250	50
stroke L [mm]	250	400
structure eigenfrq. ω_{os} [rad/s]	157	>1500
slider friction F_R [N]	60	60
maximum force F_{max} [N]	2200	2200
reluctance force F_{rel} [N]	10	10
m_1 [kg]	142	67
m_2 [kg]	125	-
electrical time lag T_{elr} [s]	$2 \cdot 10^{-4}$	$2 \cdot 10^{-4}$
delay time T_{dead} [s]	$1.5 \cdot 10^{-4}$	$1.5 \cdot 10^{-4}$

Table 4

The direct driven X-axis yields very high controller gains (see table 5), while the flexible structure causes a significant overshoot of the direct driven Y-axis, even with moderate gains. The simulated and measured tool trajectories are shown in figure 8. For the investigated machine tool, the direct drive itself will not increase the dynamic performance [2]. Further efforts to increase the structure (column) stiffness will be necessary.

controller gains	Y-axis	X-axis
position gain K_{px} [1/s]	67	67
velocity gain K_{pv} [Ns/m]	500000	10000
integrator time constant T_n [s]	0.002	0.0025

Table 5

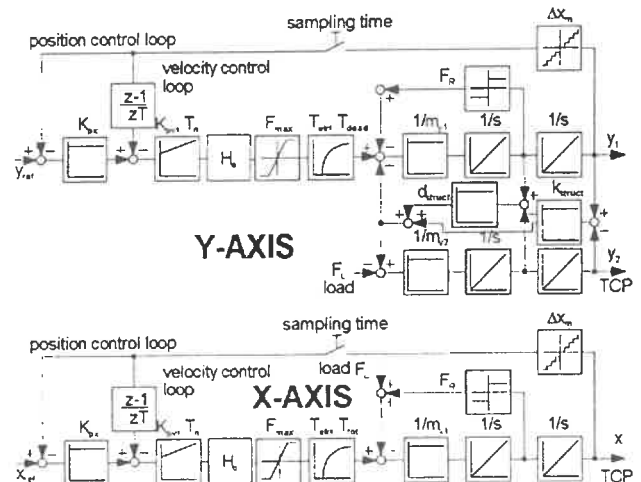


Figure 7: Block diagram with direct drives

The comparison of measurement and simulation shows good agreement also in this case. Therefore the model in figure 7 is suitable to assess the controller gains for the design and setup of direct driven machine tools.

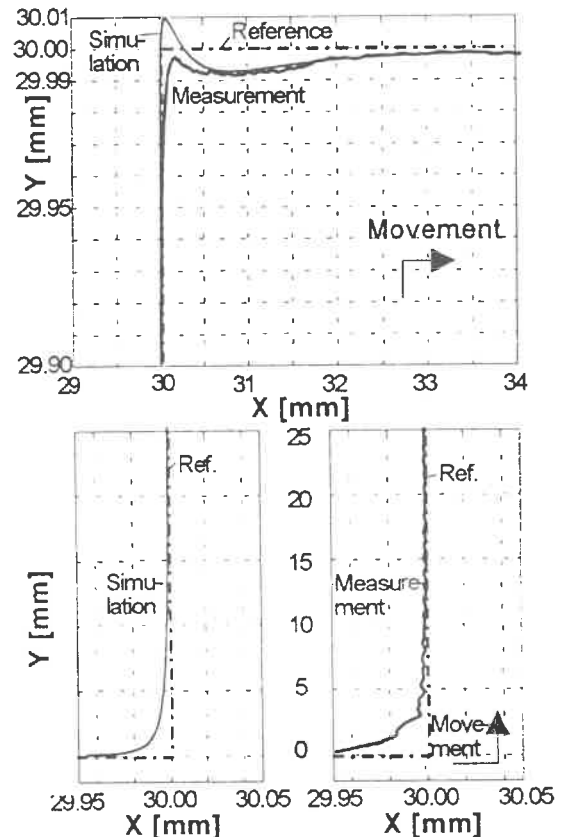


Figure 8: path monitoring of a direct driven machine

5 Performance limitation by structure flexibility

For high speed machine tools, the flexible structure becomes a very important design parameter. Figure 9 illustrates the simulation results for the position gain of direct drive and ball screw drive vs. eigenfrequency. This shows clear that stiff structures are an essential supposition for machine tools to benefit from the advantages of the direct drive.

The dotted curve in figure 9 shows the theoretical limit for the direct drive position gain depending on the structure's eigenfrequency. This represents a comprehensive design aid for the development of high dynamic machine tools.

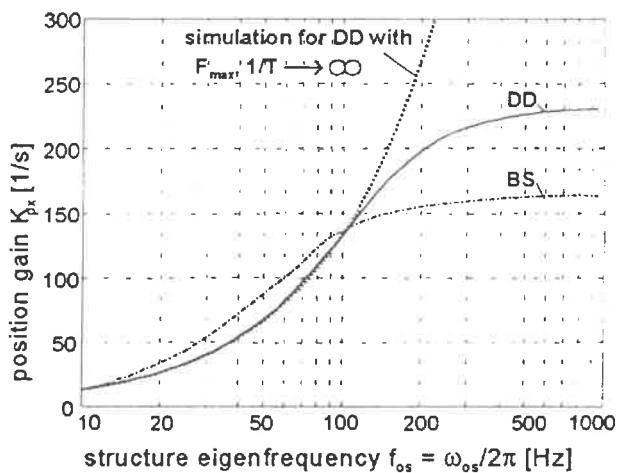


Figure 9: Position gain vs. eigenfrequency of direct drive (DD) and ball screw drive (BS)

6 Conclusion

The 2-D-measurement grid in combination with the presented model of the machine tool permits a fast monitoring and optimization of the machine tool performance. The exemplary investigation of a milling and engraving machine shows clear that the possibilities of tuning the performance of serial kinematics by using direct drives are restricted by reason of the high requirements on structural stiffness. Therefore the recently developed parallel kinematics (e.g. Hexaglide etc.) with small masses, stiff structures (struts) and high eigenfrequencies are the suitable machine principles for highest requirements on dynamic machine performance.

Acknowledgments

This project was supported by the Swiss Commission for the Promotion of Applied Research (KTI) and by the partners Heidenhain (CH) AG, ALMAC SA, ETEL SA, MIKRON AG.

Literature

- [1] Bianchi, G.; Paolucci, F.; Van den Braembussche, P.; Jovane, F.: *Toward Virtual Engineering in Machine-Tool Design*. Annals of the CIRP Vol. 45/1/1996: 379-382.
- [2] Zirn, O.; Weikert, S.; Rehsteiner, F.: *Design and Optimization of Fast Axis Feed Drives Using Nonlinear Stability Analysis*. Annals of the CIRP Vol. 45/1/1996: 363-366.
- [3] Zirn, O.; Weikert, S.; Rehsteiner, F.: *Design and Optimization of Fast Axis Feed Drives Using Nonlinear Stability Analysis*. Annals of the CIRP Vol. 45/1/1996: 363-366.

Design and Characterisation of Nanometer Precision Mechanisms

(Paper II)

Y. Xu, P. D. Atherton, P. Rhead, T. R. Hicks, and M. McConnell
Queensgate Instruments Ltd. Waterside Park, Bracknell, RG12 1RB, UK

To meet the ever increasing precision required of positioning and scanning devices, Queensgate Instruments is developing a NanoPositioning technology, the technology of moving and measuring with subnanometer precision, which helps researchers and manufacturers in the production of some of the most advanced products in use today. NanoPositioning products are developed by combining Queensgate's piezoelectric and NanoSensor technology into multi-axis positioners with the ability to position with sub-nanometer accuracy. The positioners are composed of a series of stages, called NanoMechanisms, including single axis stages, x-y stages, tilting stages and yaw stages etc.. The combinations of these stages can provide three, four, five or six degree of freedom positioning units. The nano-mechanisms are all driven by piezoelectric actuators and their positions are monitored by capacitance sensors which are of noise level $< 0.005 \text{ nmHz}^{-1/2}$ and linearity $< 0.01\%$. The outputs of the sensors are used as feedback signals to form closed loop controls. Each stage consists of a linear flexure system machined from a single Super Invar block using EDM. The NPS 3000 controller system features the use of Digital Signal Processor technology to produce a high performance servo loop, with up to 21 bit resolution. Advanced digital algorithms reduce the effect of mechanical resonances greatly improving settle times. The working bandwidth can be controlled by computer and loop parameters be defined by the user for performance optimization.

The design philosophy and some technique considerations in the development of NanoMechanism have been introduced in the previous papers^{1,2}. In this paper, the characterisation of the positioning systems is discussed and the performance details measured from these NanoMechanisms are presented. To know what the NanoMechanisms are really doing at the subnanometer level, systematic characterization has to be carried out using advanced metrological technologies. The specifications have to be carefully defined and the measuring methods be chosen properly.

Noise measurements

As it was discussed in the paper (I)¹, the measuring and positioning resolution is directly related to the noise level of the system. For a closed loop stage system, internal sensors are used for reading the actual position of the stage and providing feedback signals to the control loop. The noise measured from the sensors is the measuring noise which can not be considered as a true representation of the actual stage position, because the measuring noise includes both sensor noise and position noise. The sensor noise could be of the same level as the positioning noise, or even higher, depending on the bandwidth with which the sensor noise is measured. To monitor stage position and provide feedback signal to the controller, the internal sensor has to have a working bandwidth much higher than the stage bandwidth, which is mainly limited by the open loop resonant frequency of the mechanism. Normally, the noise level goes up as the working bandwidth increases. Both sensor noise and piezo drive noise³ in the loop will be transferred into actual displacement by the piezo actuator driven by the controller. Usually, the sensor noise is much higher than the piezo drive noise. However it can only be transferred into actual stage displacement partially because the stage working bandwidth is much lower than sensor's working bandwidth and it does not respond to the high frequency noise. Therefore, the position noise is normally lower than measuring noise. To measure the stage noise directly requires

a sensor which should have a noise level of tens pico meter or less at a working bandwidth of a few kHz. The position noise level can be worked out indirectly by measuring the noise level from HV output of the controller. This voltage noise will be charged on piezo actuator and transferred into motion noise of the stage. By calibrating the converting factor between voltage value of HV and the stage motion, the actual position noise of the stage can be calculated.

Non-linearity, hysteresis .

Non-linearity and hysteresis are motion errors along the moving axis. To assess those static motion errors, the stages are calibrated using a Zygo ZMI 1000 interferometer combined with a Nanosensor³ (capacitance sensor). After a stage is built, it is firstly calibrated against a laser interferometer and the linearity errors are measured as shown in Figure 1 (curve *a*) (the data of Figure 1 is from NPS-X-15F). Based on the initial scanning data, the internal capacitance sensors are linearized digitally to fourth order. Then, a NanoSensor is used to measure the residuals down to subnanometers. A linearity < 0.01 % can be achieved without extra effort, as shown in Figure 1 (curve *b*). Expanding the scale of the curve *b*, a plot of hysteresis error of the system is generated, as shown in Figure 2. The residuals of curve *b* is mainly limited by hysteresis, drift and noise. Subtracting the 'down' scan from the 'up' scan gives the amplitude of the hysteresis error. For the NanoMechanisms, the hysteresis errors are normally controlled less than 0.005 %, i.e. less than 0.5 nm over 10 μm .

Parasitic motions

The parasitic rotations affect the positioning accuracy due to Abbe errors. These are measured using an auto-collimator with resolution of 0.1 μrad . For single axis stages the rotation error is normally controlled within 0.1 $\mu\text{rad}/\mu\text{m}$ about the axis of maximum error and significantly less about the other axes (< 1 μrad over whole travelling range).

In some applications the out-of-plane motion of the x-y stage is important.. The out-of-plane motion is measured using a capacitive NanoSensor. In this case, the sensor is mounted parallel to the assumed x-y plane, so that the parasitic motions in the z axis will be monitored. The working area of the target electrode has to be much larger than that of probe to avoid the influence of boundary effects when the probe moves laterally relative to the target. Figure 3 shows a two dimensional profile of the out-of-plane motion of NPS-XY-100A stage over a scanning range of 100 μm x 100 μm . The peak-to-peak residuals of the out-of-plane motion from best fit plane is measured as less than two nanometers which includes the effects of measurement noise and nonlinear drift, leading to an estimation of about one nanometer for out-of-plane motion. The out-of-plane motion of the x-y stage is mainly caused by the distortions due to driving forces. The testing results indicate that the stiffness of NPS-XY-100A stage in the z axis has been designed high enough to withstand parasitic forces.

Dynamic response

In open loop, a stage can be simply considered as a mass – spring – damping second order linear system. The response of the stage is generally dominated by the stage resonant frequency, which mainly depends on the mass of the moving part and the stiffness of the system. It has a translation function

$$G(s) = \frac{1}{\frac{1}{\omega_n^2} s^2 + \frac{1}{\omega_n Q} s + 1} \quad (1)$$

Where ω_n is the resonant frequency and Q is the amplification factor. The Q value depends on the energy loss in the system or damping. For the closed loop case, the dynamic characteristics of the system are determined by both functions of controller and mechanical properties of stage. Figure 4 shows a block diagram of a closed loop PID control model of the NanoMechanisms.

The settling time is used as an important criterion which characterises the dynamic properties of a closed loop system. The shorter the settling time, the faster the stage can move. Settling time is defined as the time taken for stage to move and settle down within a certain percentage of the command step, (usually 2 %). For a step response the settling time is composed of two parts: the time for slewing and the time for ringing to settle down. For large signals the settling time mainly depends on slew-rate, and the slewing time dominates. With small signals the ringing time dominates. The small signal settling time mainly depends on PID parameters, and mechanical damping in the stage helps greatly to improve the settling speed. In most applications, the stage is driven with a series of small step signals and so the small signal settling time is the main concern. The small signal settling time is measured through a snapshot of a 500 nm step response. The time for stage to settle down to within 10 nm (2 % of the step), or sometimes 5 nm or even 2 nm for some special applications, is taken as the settling time of the stage. Figure 5 (curve *a*) is a typical step response of the NPS-Z-15A, which has a small signal settling time of 1.6 ms for settling down to 2 % of the command position.

Sometimes, the loading mass can have a big influence on dynamic performance of the system, especially for the stages of compact size. To avoid the problem, the Queensgate digital controller is so designed that the PID parameters can be easily defined by users and therefore it possible to optimize the performance in situ.. Curve *a* in Figure 5 is measured from an unloaded stage under the control of optimized PID parameters, curve *b* shows how the performance was influenced by a 100 g mass loaded onto the stage, and re-optimized PID parameters gives back a desired step response as curve *c*.

Conclusion

To build an ultra precision positioning stage of nanometer or even subnanometer metrological capability, many metrological and design concepts have to be considered carefully. Those mainly include noise, linearity, hysteresis, Abbe errors (linear and rotational parasitic motions), dynamic characteristics and stability. The NanoMechanisms are designed with all these issues in mind and built and characterized with most advanced techniques. The initial results of systematic assessment have shown that their performance is well under control at nanometer level or below.

Reference

1. Ying Xu, Paul D Atherton, Thomas R. Hicks and Malachy McConnell, Design and Characterisation of Nanometer Precision Mechanisms (Paper I), *Proceedings of ASPE 1996 Annual Meeting*, Monterey.
2. Thomas R. Hicks, Paul D Atherton, Ying Xu, and Malachy McConnell, *The NanoPositioning Book*, Queensgate Instruments Ltd, 1997
3. NanoSensor is a trade mark for the capacitance sensors of Queensgate Instruments Ltd

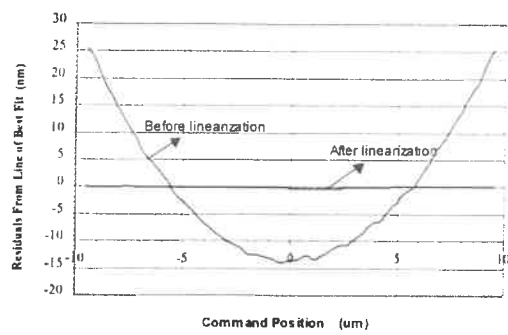


Figure 1 Linearity errors

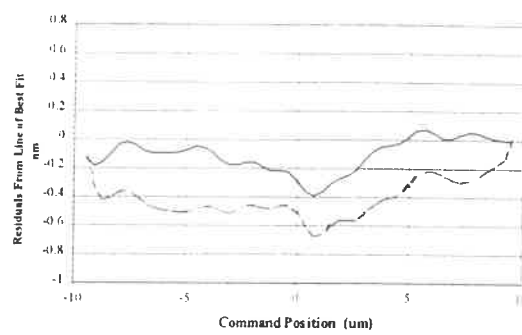


Figure 2 Hysteresis error

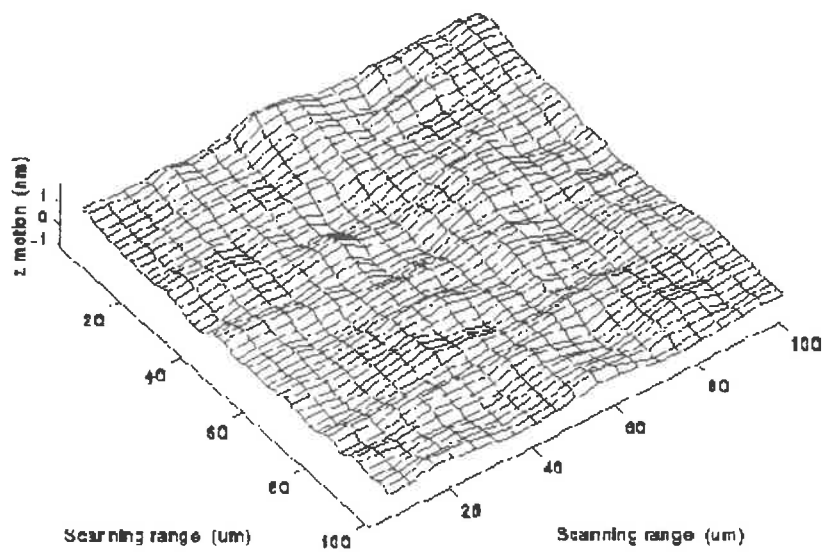


Figure 3 Out-of-plane parasitic motion of x-y stage

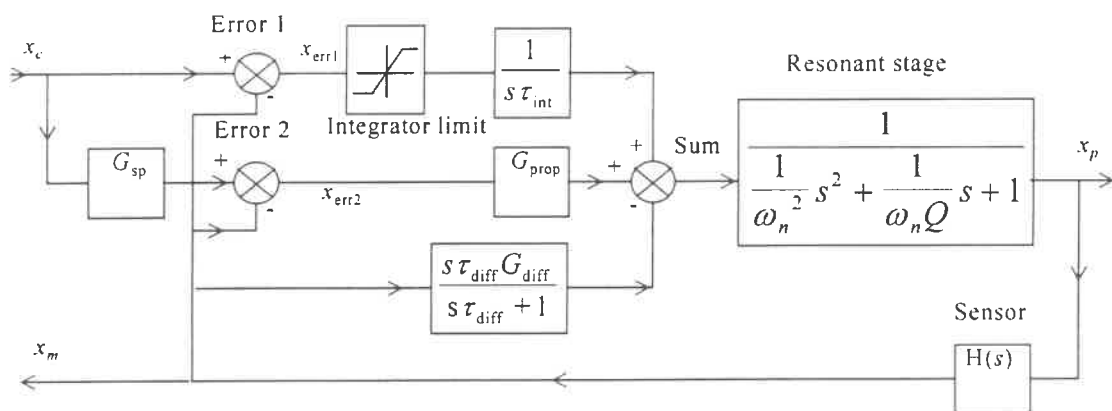


Figure 4 A block diagram of closed loop control

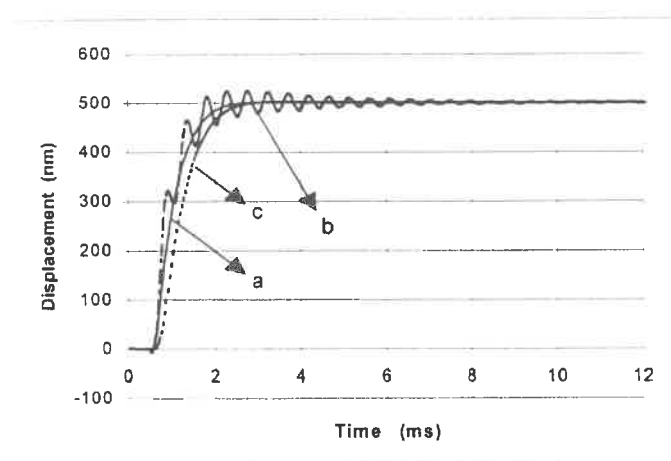


Figure 5 Step response of NPS-Z-15A

PRELOAD CONTROL IN BALL SCREW USING PIEZOELECTRIC ACTUATORS

Shigeo Fukada
Department of Mechanical Systems Engineering
Shinshu University
500 Wakasato, Nagano 380 JAPAN

1. Introduction

Ball screws are usually used in the preloaded condition. The preload serves to eliminate the backlash of nuts and to improve their stiffness. On the other hand, the preload also has negative effects: it increases friction loss, which results in more heat generation, causing thermal expansion of the screw shaft and accuracy deterioration¹⁾²⁾. Although the preload has these advantages and drawbacks, the existing ball screws used are under constant preload. Therefore, it is desirable that the preload be controlled: increased when the load is heavy and the speed is low; decreased when the load is slight and the speed is high. This study proposes a method of controlling the preload of ball screws using piezoelectric actuators (PEA). The expansion of the PEA is so slight that the stiffness of each part must be painstakingly designed to transform the expansion into the force efficiently. In this report, first, the force generating characteristics of the PEA are considered fundamentally using a simplified model; next, manufacture and testing of the preload controllable ball screw are examined.

2. Principle

The concept of the preload controllable ball screw is shown in Fig. 1. The ball screw is the double-nut type, and PEAs are inserted between the nuts. The preload is controlled by PEAs to change the distance between the double nuts. In order to eliminate the hysteretic effect of PEA, the actual preload is sensed by load cells for feedback control.

A statical model for the mechanism of preload control is described in Fig. 2. The PEA is modeled as a combination of a generator of displacement δ_p and an elastic element of stiffness K_2 . Fig. 3 shows the characteristics of the PEA used in this study. If the expansion of the PEA is not restrained, it will generate a displacement of 16 μm for the applied voltage of 100 V. On the other hand, if the expansion is completely restrained, it will generate a force of 4 kN. However, in actual practice, the displacement of PEA is absorbed

into other elastic elements, so that the generated force is reduced.

In the model of Fig. 2, the generated force F is given by the following equation.

$$F = K_0 (\delta_0 + \delta_p + \delta_d) \quad (1)$$

where δ_0 is the displacement of initial preloading, and δ_d represents the change in the distance between the

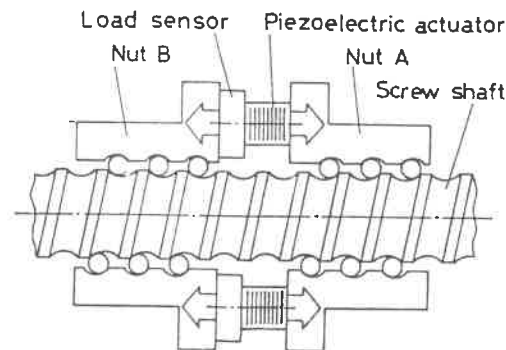


Fig. 1 Basic concept of preload control

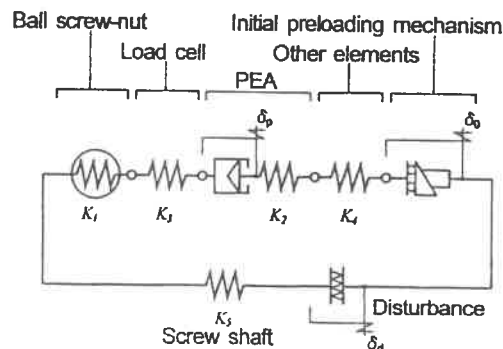


Fig. 2 Principle of preload control

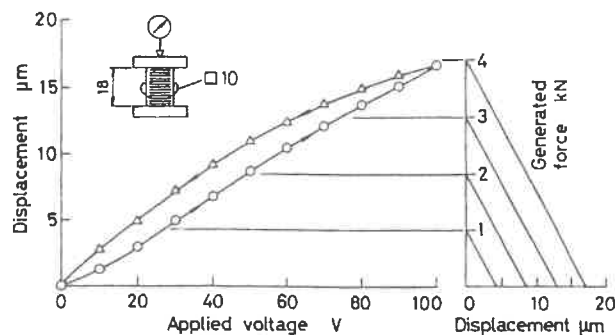


Fig. 3 Characteristics of piezoelectric actuator

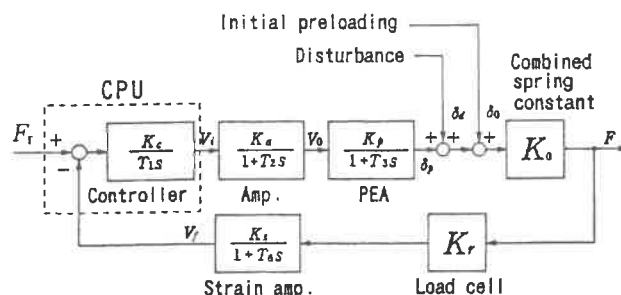
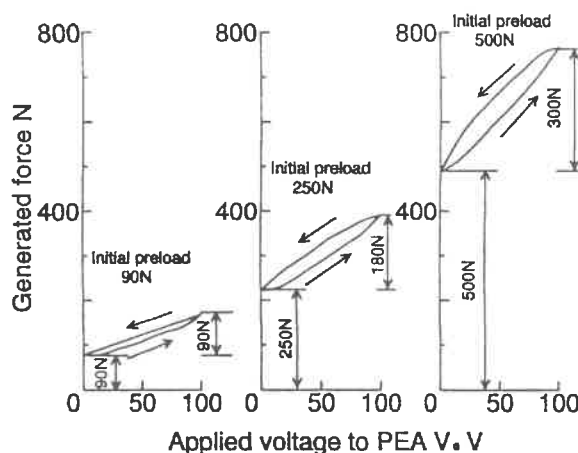
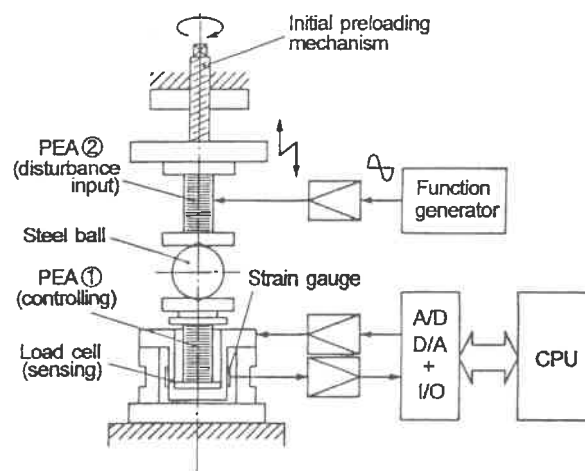
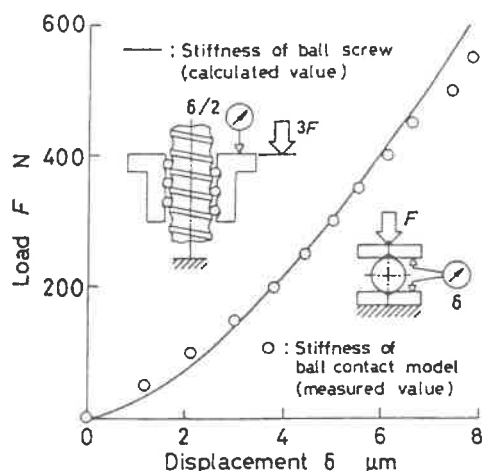


Fig. 7 Block diagram of control system

3. Experiments on simplified model

3.1 Static characteristics First, a compact preloading unit, which consists of a PEA ($\square 10 \times 18$ mm) and a cylindrical load cell with strain gauges, is manufactured for trial. To evaluate the preloading unit, simplified experiments are performed as shown in Fig. 5. The stiffness of the practical ball screw is simulated by a "single ball model". The measured stiffness of the model is similar to the ball screw as shown in Fig. 4. To simulate the disturbance in Fig. 2, another PEA is set serially to the preload control unit.

Fig. 6 shows the generated force as the function of applied voltage to the PEA. The force at this point is not controlled with feedback, so that the characteristics show hysteretic effects. It is shown in the figure that the manipulating range of the preload is enlarged by the increase of the initial preload because of the nonlinearity of the stiffness mentioned above. The range of preload control is about 300 N with an initial preload of 500 N.

nuts caused by the lead error of the screw shaft, which acts as the disturbance for the preload control system. K_0 in eq. (1) means the combined stiffness of the mechanism, given as follows.

$$\frac{1}{K_0} = \sum_{i=1}^5 \frac{1}{K_i} \quad (2)$$

K_1 in Fig. 2 represents the stiffness between the screw shaft and the nuts through circulating balls. Because the characteristics of K_1 are ruled by Herzian contact, it is nonlinear as shown in Fig. 4. The solid line in the figure represents the stiffness of the ball screw used in this study³⁾. The stiffness, as seen in the slope of the curve, becomes larger as the applied load increases, so that the manipulated range of the preload becomes wider with the increase of the initial preload.

3.2 Control system The generated force is sensed by the load cell, and the sensed signal is fed back to the controller. Fig. 7 shows a block diagram of the control system. The controller with integral operation is realized by 16 bits digital CPU. The sampling time T_1 is 2 ms. The parameter K_c is tuned to 1.4 V/V.

Fig. 8 shows the closed loop response of the control system to the step input of reference F_r . The resolution of preload control is finer than 5 N, and the system responds without hysteresis.

3.3 Dynamic response Fig. 9 shows the frequency response of the closed loop system: (a) is the response to the reference input. The band width of the system is limited to 10 Hz even though the PEA must have a dynamic range wider than 10 kHz. The cause of the limitation lies in the low pass filter in the strain amplifier, the cut off frequency of which is 10 Hz.

Fig. 9 (b) is the response to the disturbance input. The disturbance is given by applying a sinusoidal voltage of 12 V amplitude to the PEA② in Fig. 5. The preload control system responds to the dynamic input to resume the disturbance of 5 Hz.

4. Preload controllable ball screw

4.1 Structure On the basis of the results of the above simplified experiments, a preload controllable ball screw is constructed as shown in Fig. 10. Three preloading units are set on the nut flange so as to surround the screw shaft symmetrically; then they are inserted between the double nuts of the ball screw. The outer diameter of the screw shaft is 32 mm, the lead is 5 mm, and the circuit of ball circulation is 2.5×2 . Micrometer heads are used as the initial preloading

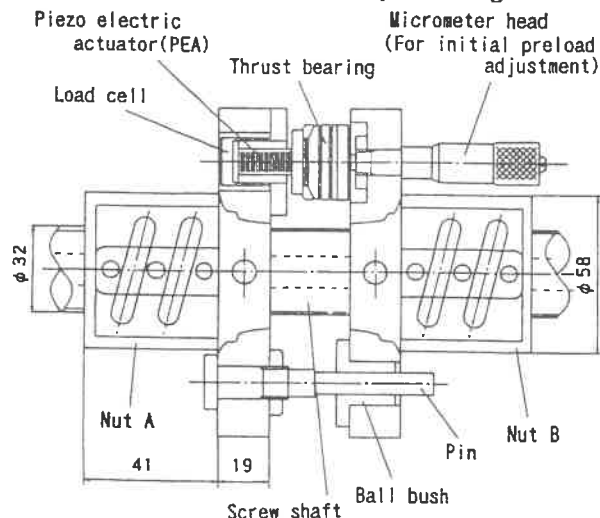


Fig. 10 Construction of preload controllable ball screw

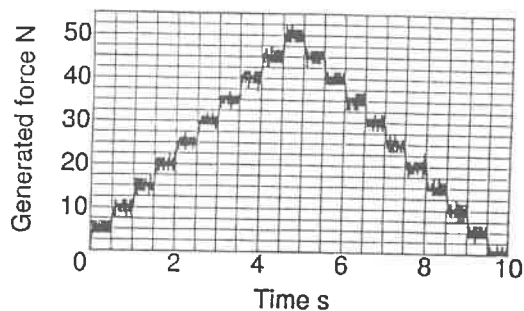
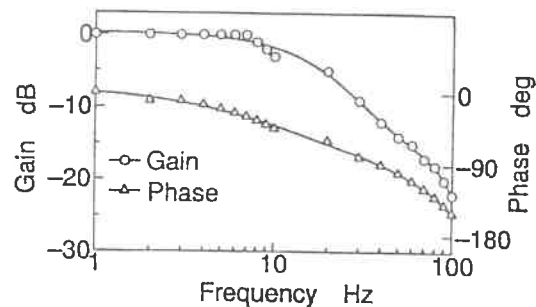
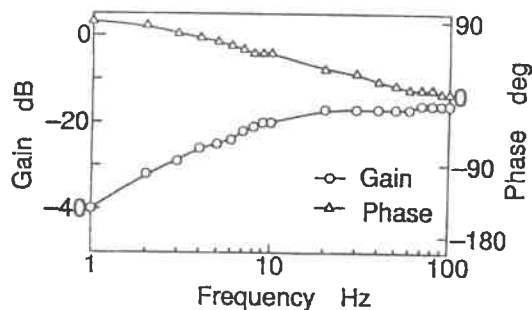


Fig. 8 Step response of preload control system

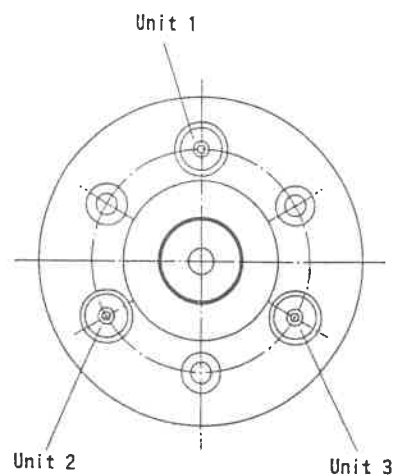


(a) Response to reference input



(b) Response to disturbance input

Fig. 9 Frequency response



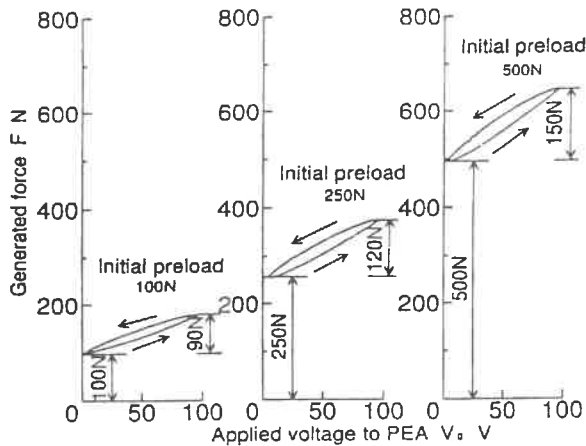
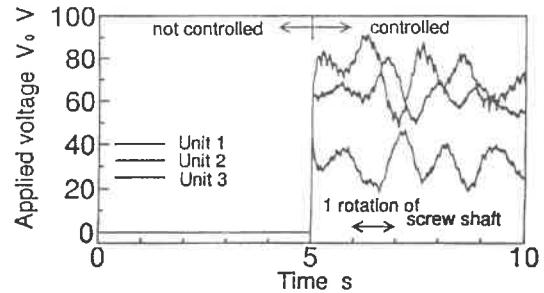


Fig. 11 Initial preload and force generation

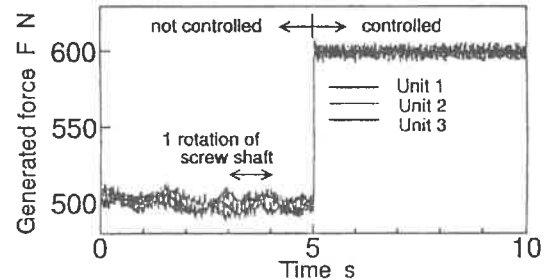
device. A thrust ball bearing with spherical washer is inserted between each micrometer head and the preload unit so as to prevent the misalignment. To fix the phase between the nuts, three pins guided by ball bushes are inserted through the nut flanges.

4.2 Performance First, the range of generated force is tested without feedback control, similar to the experiments of Fig. 6. Fig. 11 shows the generated force per preload control unit as the function of the applied voltage to the PEA. The range of preload control is 150 N with an initial preload of 500 N, or a total range of 450 N for three units. The actual range per unit is smaller than the result on the simplified model. The actual mechanism in Fig. 10 includes some elements with low stiffness, such as the thrust ball bearings or micrometer heads, which are not assumed in the simplified model, so that the combined stiffness K_0 in eq. (1) is smaller than in the simplified model.

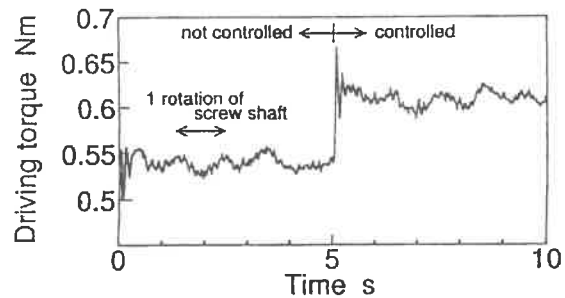
Next, the preload is controlled with feedback control system as shown in Fig. 7. The experiments are performed on the ball screw while it is rotating with an actual speed of 60 rpm, while the driving torque (drag torque) is measured instantaneously. Fig. 12 shows the results. (a) shows the applied voltage to the PEAs, and (b) shows the generated force of each unit. The initial preload is 500 N per unit, and the desired value of preload is 600 N. After feedback control starts, the applied voltages are regulated so as to keep the preload at 600 N. Variation of the preload caused by rotation is eliminated by the preload control. Fig. 12 (c) shows the driving torque of the ball screw. The torque changes at the same time as the preload is increased. These results mean that preload control in ball screw is possible using PEA, and that this preload control is available to change the driving condition dynamically or actively.



(a) Applied voltage



(b) Generated force



(c) Driving torque

Fig. 12 Performance of preload control

5. Conclusions

- (1) A method of controlling the preload of ball screw is proposed using PEAs, and a compact preloading unit is manufactured for trial.
- (2) Experiments on the simplified model show that the control range of the preload is about 300 N with an initial preload of 500 N.
- (3) The preload controllable ball screw is constructed using the preloading units, and the availability of the preload control is confirmed experimentally.

References

- 1) J. Otsuka, S. Fukada et al.: A Study of Thermal Expansion of Ball Screw, Bull. JSPE, 17, 4 (1983) 273.
- 2) N. Obuchi and J. Otsuka: Study of Friction Torque of Ball Screw, Bull. P.M.E. (Tokyo Inst. Tech.), No. 61(1988)1.
- 3) NSK Ball Screw Catalog, CAT. No. 3101 (1993).

Thin, Light-weight, Compact Pick-Up Actuator for High Precision, Fast Access CD-ROM Devices

S. M. KIM, D. G. GWEON

*Dept. of Mechanical Engineering
Korea Advanced Institute of Science and Technology*

Abstract

For a conventional pick-up actuator, the magnets are used separately to the coarse actuator and the two axis actuators. That is, a magnet and yoke assembly is mounted on the moving part which becomes heavy. But in this study the proposed thin, light-weight, compact pick-up actuator is designed to mount only one magnet and yoke assembly for the coarse actuator and two axis actuators to reduce the moving part mass, and the teflon coating was done to the guide to reduce friction, which results in the significant decrease in seek time. For the optimally designed pick-up actuator, the thickness and the seek time was obtained below 10mm and 15msec, respectively. The parasitic resonances due to the wire suspension supporting the optical head are attenuated using the modal strain energy method.

1. Introduction

The small sized optical disk drive with short seek time capability is generally applied to the multi-media devices and computers. To reduce the seek time, the optimal design study for the moving coil actuator was being done.^{[1],[2]} However, the optimal design was mainly focused to the coarse actuator. The usual pick-up actuator used the magnet separately on the coarse actuator and two axis actuators. That is, a magnet and yoke assembly are mounted on the moving part, which leaved the moving part heavy. But in this study the proposed thin, light-weight, compact pick-up actuator is designed to mount only one magnet and yoke assembly on the coarse actuator and two axis actuators to reduce the moving part mass, and the teflon coating was done to the guide to reduce friction. The moving coil actuator is designed via the optimal design including the coarse actuator and two axis actuators.

2. Structure of the moving coil actuator

The proposed pick-up actuator is shown in Fig.1. As can be seen from the Fig.1, a fixed magnet and yoke assembly is used for the coarse and two-axis actuators. For the coarse actuator, teflon coating was processed on the guide to reduce friction. The optical head is supported by four wire suspensions. From this optimal design, the pick-up actuator can be made with less than 10mm thickness and light weight.

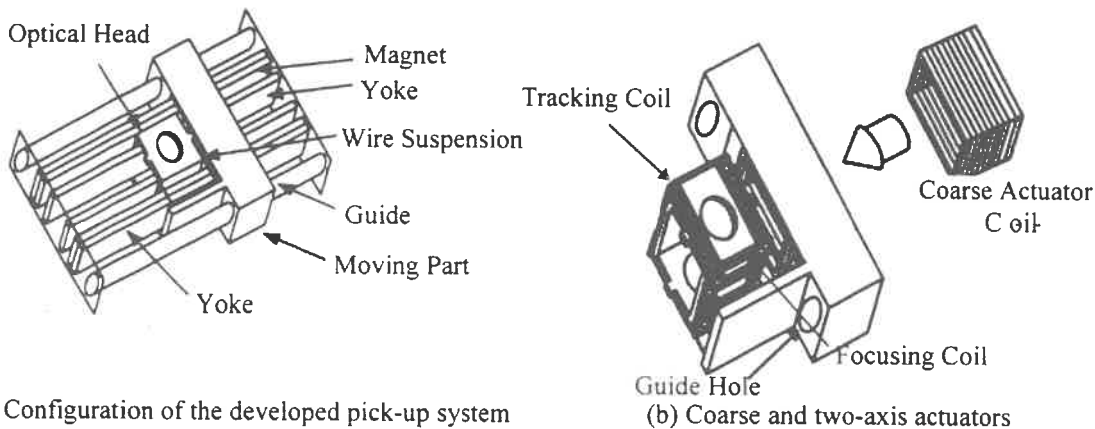


Fig. 1 Structure of the pick-up actuator

3. Optimum design of the pick-up actuator for fast access

The optimal design parameters of pick-up actuator are the dimensions of moving part, the coarse actuator coil, the focusing/tracking coil, the optical head, the magnetic flux density. The cost function in the optimal design is to minimize the mass of the optical head and the moving part. The constraints of the optimal design are the 1st natural frequency (higher than 25Hz), 1Hz sensitivity (1mm/V), 200Hz sensitivity (50 μ m/V) and Q-value (lower than 8dB).

Hence, the proposed pick-up actuator is optimally designed to satisfy the constraints minimizing the cost function. To get 15msec average seek time⁽³⁾(1/3 full stroke), the coarse actuator drive force F_s and the required force F_n are given as below.

$$F_s = B L n I \quad (1)$$

$$F_n = M_t \left[L_s \left(\frac{\tau_a}{T_{av}} \right)^2 + \mu(v)g \right] \quad (2)$$

where

B : magnetic flux density at air gaps L : coil length in the air gaps
 n : number of coarse actuator coil turns I : coarse actuator coil current
 M_t : total mass L_s : access full stroke
 T_{av} : average seek time (1/3 full stroke) g : gravitational acceleration

$$\tau_a = \frac{\beta_m + \frac{1}{3}}{\sqrt{\beta_m}} \quad \beta_m : \text{minimum speed factor (from } \frac{\partial F_n}{\partial \beta} = 0 \text{)}$$

$\mu(v)$: friction coefficient of teflon coating for velocity

From the equation(2) the minimum value of F_n is calculated for the given average seek time. The actuator is optimally designed so that F_s has the minimum value of F_n , which is ultimately the same approach to minimize the actuator size. The total mass M_t from the equation(1) covers the coil mass determined by the number of coil turns and coil length in equation(1). M_t also includes the optical head mass. Therefore it is impossible to obtain the F_n and F_s directly. That is, F_s can be calculated numerically. Figure 2 shows the flow chart in the optimization program.

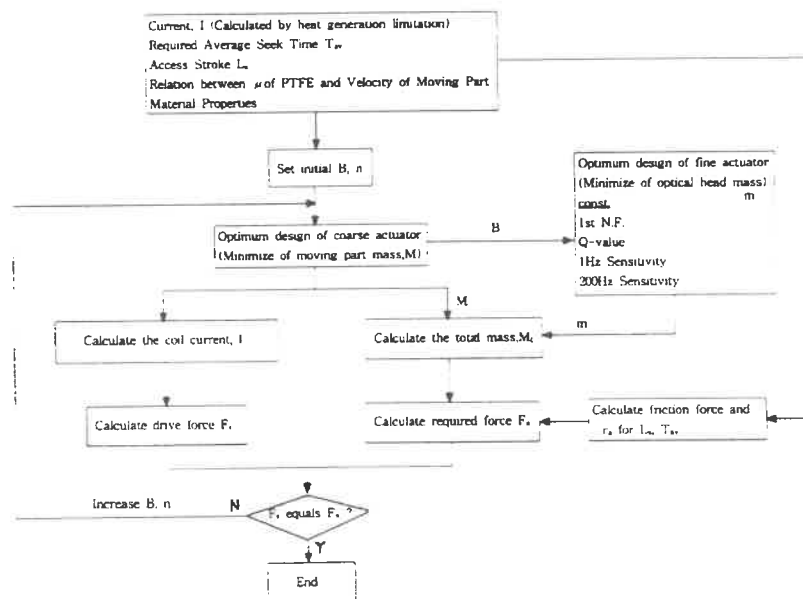


Fig. 2 Flow chart for the developed actuator design

From the fig.2, the initial small values of B , n are arbitrarily selected. The design parameters and F_s are calculated satisfying the cost function and constraints. If F_s is not equal to F_n , iterations are repeated so that F_s is converged to F_n by increasing the infinitesimal amount to B and n . Figure 3 shows the relationship between the speed factor and average seek time with respect to the M_t of 15 g and L_s of 32 mm. As can be seen, the minimum average seek time is obtained when β value is 0.16.

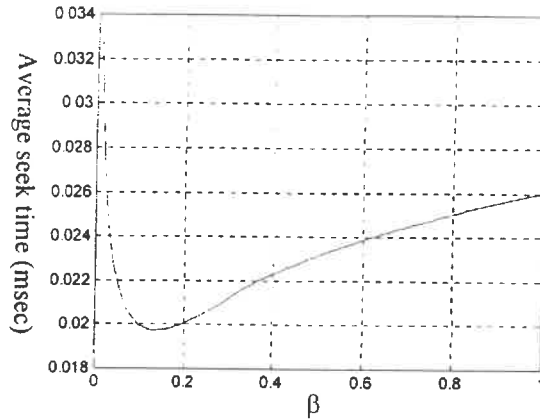


Fig. 3 Average seek time vs. speed factor

4. Vibration and damping analysis of wire suspension

From the fig.4 assuming a quarter of the optical head mass is supported by each wire suspension, the parasitic resonances of wire suspension is attenuated by the viscous damping material.

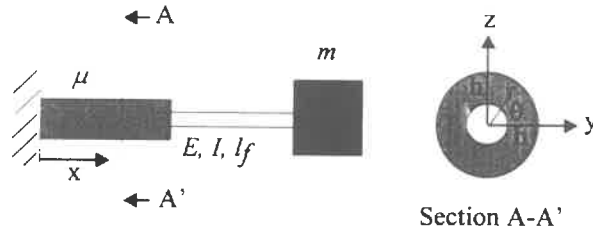


Fig.4 A wire spring modelled as cantilever beam

In the Fig. E , I , and l_f are representing the elastic modulus, moment of inertia, length of the suspension, while m and μ represent a quarter of optical head mass and the dynamic viscosity of viscous damping material, respectively. Also, h represents the gap between the wire suspension and the damping material. The damping ratio can be obtained using the modal strain energy method^[4]. The beam's transverse displacement with respect to each mode can be expressed as follows.

$$Z_i(x, t) = Z_i(x) e^{j\omega_i t} \quad (3)$$

The beam's damping ratio in each mode can be written as equation(4).

$$C_i = \frac{1}{2} \left(\frac{W_{diss}}{2\pi W_{total}} \right) \quad (4)$$

where W_{diss} represents the dissipated energy by the damping material during one cycle vibration while W_{total} represents the total energy stored to the beam during resonance. Using the equation (3), W_{total} can be calculated as given in equation (5).

$$W_{total} = \left| \frac{1}{2} \rho \int_0^{l_f} \left(\frac{\partial Z_i}{\partial t} \right)^2 dx \right| \quad (5)$$

where $l_f \rho$ represents the mass of wire suspension. Using the transverse displacement, $Z_i(x, t)$, beam's longitudinal displacement, $\delta(x, t)$, can be obtained as below.

$$\frac{\partial \delta_i(x, t)}{\partial x} = -r \sin(\theta) \frac{\partial^2 Z_i(x, t)}{\partial x^2} \quad (6)$$

Assuming the dissipated energy is caused by the shear deformation of damping material which appears to the surface of wire suspension, the dissipated power is written as follows.

$$\Pi_{diss} = \frac{\mu\omega r}{l_f} \left| \int_A \delta_i \dot{\delta}_i d\theta dx \right| \quad (7)$$

From the equation (7), the dissipated energy can be obtained as below.

$$W_{diss} = \left| \Pi_{diss} \right| \frac{\pi}{\omega r} \quad (8)$$

Hence the damping ratio can be calculated using the equations(4), (5), and (8). The vibration and damping analysis of the wire suspension using the modal strain energy method and the optimization algorithm in fig.2 are used to design optimal pick-up actuator.

Table 1 lists the performances of the optimal-designed actuator when T_{av} is 15msec. The performances listed in the Table 1 satisfy all the optimization constraints.

Table 1

1 st natural frequency (tracking)	27.74 Hz	Q-value (focusing)	6.16 dB
Q-value (tracking)	8 dB	1Hz sensitivity (focusing)	6.47 mm/V
1Hz sensitivity (tracking)	5.72 mm/V	200Hz sensitivity (focusing)	126.6 μ m/V
200Hz sensitivity (tracking)	112.2 μ m/V	Magnetic flux density (B)	5100 Gauss
1 st natural frequency (focusing)	27.74 Hz	Total mass (M _t)	1.48 g

Figures 5 and 6 show the frequency responses to the direction of focusing direction with and without damping, respectively. In fig.6 many parasitic resonances are happened within the 30 kHz range. However, these parasitic resonances are fully attenuated due to the damping ratio using the modal strain energy method, as can be seen from the fig.6.

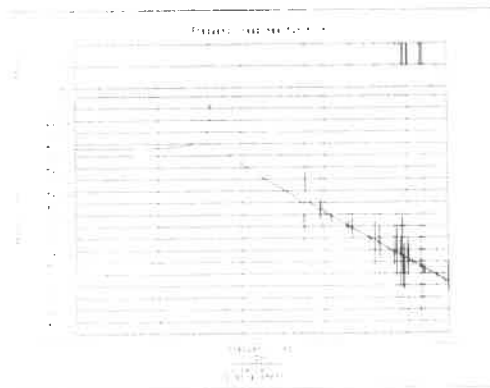


Fig.5 Frequency response without damping

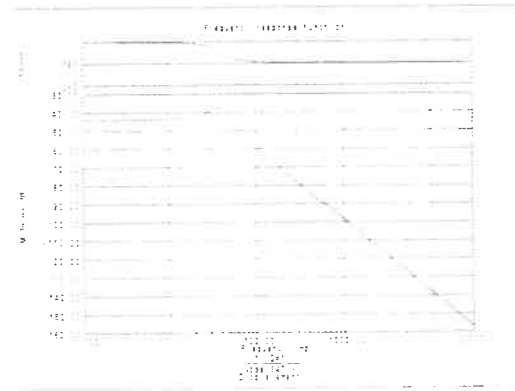


Fig.6 Frequency response with damping

5. Conclusions

For the proposed pick-up actuator in this study a fixed magnet and yoke assembly are used to the coarse and two axis actuators. This structure makes it possible to design the thin, light-weight, compact pick-up actuator with thickness of less than 10mm. Using the optimal designed pick-up actuator fast seek can be possible that the average seek time is less than 15msec. The parasitic resonances due to the wire suspension supporting the optical head are attenuated using the modal strain energy method.

REFERENCES

- [1] Tetsu Yamamoto, Takashi Yumura and Hiroo Shimegi, *Proc. SPIE*, Vol.695, pp.153-159(1986)
- [2] J. Arthur Wagner, *IEEE Trans. on Magnetics*, Vol.Mag-19, No.5, pp.1686-1688(1983)
- [3] F. R. Hertrich, *IBM J. Res. Develop.*, Vol.9, pp.124-133(1965)
- [4] Marsh, E. R. and Slocum, A. H., *J. of ASPE*, Vol.18, No.2/3, pp.103-109(1996)

A Long Stroke and Ultra Precision XY θ Stage Using Simple Linear DC servo Actuators

Sung-Q Lee and Dae-Gab Gweon

Dept. of Mechanical Engineering

Korea Advanced Institute of Science and Technology

373-1 Kusung-Dong, Yusong-Gu, Taejon, 305-701, Korea

Abstract

A XY θ stage which has a long stroke and ultra precision resolution without extra fine servo is proposed. The characteristics of the stage are symmetry, simple servo principle, and inherently infinite resolution. it has three linear actuators and three air bearings which are radially disposed with 120 degrees pitch, so that it has symmetry. And a linear actuator is the type of VCM, so that it has simple servo mechanism. Moreover, with the use of air bearings, it has no friction or wear allowing the chance of infinite resolution.

1. Introduction

A dual servo stage (coarse servo and fine servo) has been popular for the wafer positioning purpose. However, this has complex structure and control scheme. Therefore, a XY θ stage which has a long stroke and ultra precision resolution without extra fine servo - more over rapid response - is proposed. Recently, many kinds of systems are published to meet the purpose of long stroke and ultra precision resolution-for example friction drive, dual servo system in which coarse/fine servo work together and surface motor using BLDC type motor. However, though these systems meet the purpose, a friction drive has low bandwidth, a dual servo system has complex structure and control algorithms, and a surface motor has complicated motor. Therefore, a Long Stroke and Ultra Precision XY θ Stage Using Simple Linear DC servo Actuators is presented.

2. Characteristics of System

Symmetry

Figure 1. shows the configuration and prototype photo of the proposed system. The stage consists of three linear actuators and three air bearings. Three linear actuators are attached at the upper plate symmetrically and disposed radially with 120 degrees pitch, and so does the air bearings. Actuation force is induced by the moving coil type voice coil motor (VCM). From the combination of three actuators, x, y and θ motions can be obtained. Especially, the stage will be rotated when each actuator produces the same force. With this configuration, working space becomes the shape of 10mm diameter circle. By virtue of symmetrical configuration, this system is under low thermal effects and appropriate to the precision system.

Simple Servo

Figure 2. shows the cross section of the linear actuator. Actuation force is induced by moving coil type voice coil motor (VCM) as mentioned above. Since moving coil type VCM has lighter moving mass than that of moving magnet type, the moving coil type is preferable. Each actuator has permanent magnets which produce the Z-directional magnetic flux and a moving coil which cross the permanent magnets. As in figure 2, two permanent magnets are placed to look each other in order to make the same directional force when the closed coil has different current direction. By the Fleming's Left Hand Rule, each actuator simply produce actuation force.

$$F_{act} = nBil$$

where n is the number of coil turns, B , air gap flux, i , current, l , coil effective length. It is noticeable that since there are no ripples or nonlinearity of magnetic flux, high precision position control can be possible as well as high speed. By the way, like usual rotary motor, since it has also back emf, it need care to derive the transfer function from voltage input to output speed.

In this system, 2BB ferrite permanent magnet of maximum residual flux is 4,100 gauss is used. Therefore at operating point flux is about 3500 gauss. Coil of 0.5 diameter is wound 480 turns. l , effective length is 50mm. Therefore, one actuator of this system can exert following actuator force at rated 1 ampere(A),

$$\begin{aligned} F_{act} &= 480(\text{turns}) \times 0.35(\text{T}) \times 1(\text{A}) \times 0.05(\text{m}) \times 2(\text{two sides}) \\ &= 16.8 (\text{N}) \end{aligned}$$

The specifications of proposed system are summarized in Table 1

Total range	10mm diameter circle
Linear Actuator	
Permanent Magnet	
Quantity	6 EA(2EA/1actuator)
Material	Ferrite type, 3500 gauss
Coil	
Number of turns	480
Size	0.5mm diameter
Rated Current	1 ampere
Rated Actuation Force	16 N
Air gap	7mm
Air bearing	
Type	Annular type (Ro=30mm, Ri=12mm)

Table 1. Specification of Proposed System

Inherently infinite resolution

Three air bearings are used to guide the stage. Each bearing is attached to the upper plate and supports the moving part. The stage can move XY-direction freely and is under constraint in Z-direction to reduce the Z-direction movement and tilting. Because of air bearing, the stage does not have friction and wear, so that ultra precision control can be possible. Also, like other servo motors, this system has inherently infinite resolution but is restricted by the measurement

device or digital quantization. Therefore, without extra fine servo, structure and control scheme is simple and ultra precision resolution can be obtained.

3. Dynamics of this system

Figure 3 shows the schematic of three directional actuation mechanism. Each linear actuator exert actuation force are same and can be obtained by the following equation

$$F_{act} = nBil = k_s i$$

where trust constant $k_s = nBl$.

The dynamics of the system can be expressed by the following matrix equations:

$$M \ddot{x} = T k i$$

$$L \dot{i} + R i = V - k T^T \dot{x}$$

$$T = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$$

Where $M = \text{diag}(M, M, J)$ is the mass and inertia matrix, $x = [x, y, \theta]^T$ stage position vector, $k = \text{diag}(k_s, k_s, k_s)$ the trust constant matrix, $L = \text{diag}(L, L, L)$ the coil inductance matrix, $R = \text{diag}(R, R, R)$ the coil resistance matrix, $i = [i_x, i_y, i_\theta]$, the current vector, $V = [v_x, v_y, v_\theta]$ the input voltage vector, T the transformation matrix linearized at $\theta=0$ degree.

It is noticeable that the k, L, R , matrix has the value because of the symmetrical design of the system.

4. Conclusion

A XY θ stage which has long stroke and ultra precision resolution without extra fine servo is proposed. The characteristics of stage are symmetry, simple servo principle, and inherently infinite resolution. In detail, it has three linear actuators and three air bearings which are radially disposed with 120 degrees pitch, so that it has symmetry. And a linear actuator is the type of VCM, so that simple servo mechanism can be obtained. Moreover, with the use of air bearings it has no friction or wear. To control the system it need to derive the dynamics of proposed system and is derived. In the future, closed loop control may be performed with laser interferometer feedback.

References

- [1] Alexander H. Slocum, 1992, "Precision Machine Design", Prentice-Hall.
- [2] S.A.Nasar and L.E.Unnewehr, 1983, "Electromechanics and Electric Machines", John Wiley and Sons.
- [3] Yuichi Okazaki, Shin Asano and Takayuki Goto, 1991, "Dual-servo Mechanical stage for continuous positioning" Int. JSPE Vol.27, No.2, pp172-173.
- [4] Rodney Kendall, Sam Doran and Erwin Weissmann, 1991, "A servo Guided X-Y-theta Stage for Electron beam lithography", J. Vac. Sci. Technol. B9(6), Nov/Dec.
- [5] Yoshiyuki Tomita, Yasushi Koyanagawa and Fumiaki Satoh, 1996, "A Surface Motor-driven Precise Positioning System", Precision Engineering, Vol.16, No.3, pp184-191.

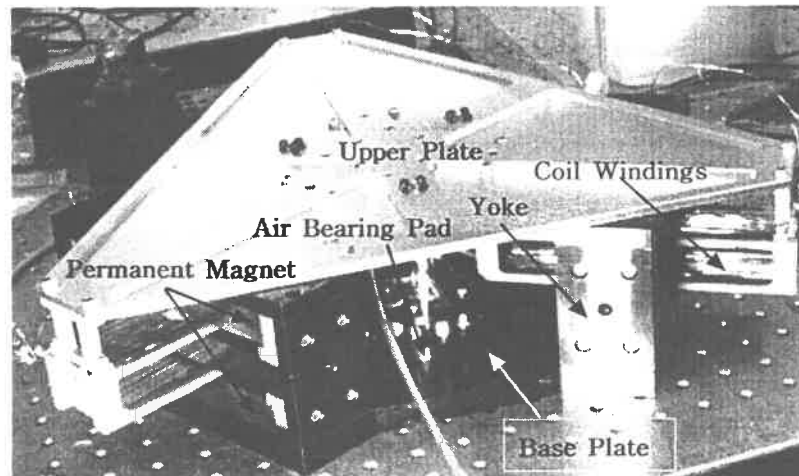


Fig. 1 Prototype of Proposed System

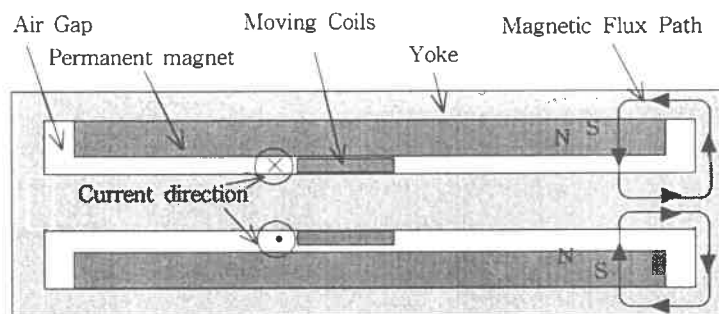


Fig. 2 Cross Section of Linear Actuator

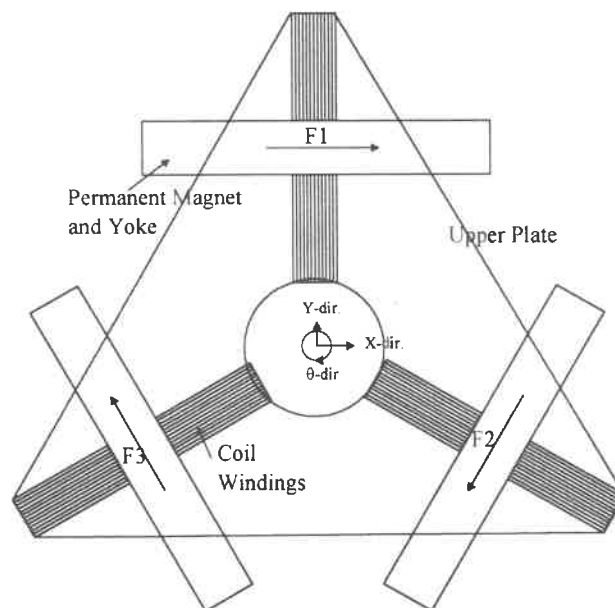


Fig. 3 Principle of Three Actuator Combination

Function Related Tolerance Theory and Its Application to Needle-Bar Motion Mechanism in A Sewing Machine

Keyoung Jin Chun, Dae Young Shin

Korea Institute of Industrial Technology
San 17-1, HongChonRi, IbJangMyun,
ChonAnSi, Chung Nam 330-820
KOREA

1. Introduction

In order to apply a tolerance theory to needle-bar motion mechanism (NMM) of an industrial lock stitch sewing machine, kinematic and sensitivity analysis are carried out. From a kinematic analysis, three constraint equations are obtained, which is later used to derive a mathematical expression of the sensitivity coefficient. Furthermore, a locus of a needle's extremity is obtained, and a dominant factor ('blade point') for stitch making is found. From a sensitivity analysis, the sensitivity coefficient of each design variable is represented as a mathematical expression by use of the implicit differentiation. Thus, the exact value of each sensitivity coefficient can be obtained. Then, it is discussed that how design variables have effects on this dominant factor, which interrelations between each design variable exist, and how tolerance values are effectively allocated to design variables in

consideration of machining and/or assembling difficulties within the domain of allowable errors of the machine's function.

From the above results, the motion of NMM is characterized, and well distributed tolerance values to design variables might be determined by a function related tolerance theory for better machine designs.

2. Kinematic analysis of NMM

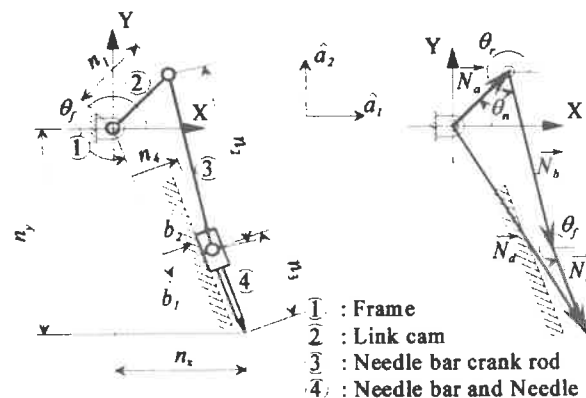


Figure 1. Needle bar motion mechanism

NMM is a slide crank mechanism as shown in Figure 1. The reference system is the Cartesian coordinate system. There are two unit coordinate vectors on a frame.

One is $(\hat{a}_1, \hat{a}_2)$ directed along the X and Y coordinate axes. The other is $(\hat{b}_1, \hat{b}_2)$ which defines the sliding plane S of a follower. Here, $\hat{b}_1$ is parallel to the sliding plane S.

NMM is composed of the link cam (driver), the needle bar crank rod (coupler), the needle bar and needle (follower) and the body (frame) as shown in Figure 1. The input of NMM is a rotational angle θ_n of the link cam and the output of NMM is a periodic movement (n_x, n_y) of the needle's extremity. NMM has five constants $(n_1, n_2, n_3, n_4, \theta_f)$, one independent variable θ_n and three dependent variables θ_r, n_x, n_y . Here, θ_f and n_4 define the sliding plane of the follower.

There is a closed loop $(\bar{N}_a + \bar{N}_b + \bar{N}_c - \bar{N}_d = 0)$ as shown in Figure 1. As Cartesian components on $\hat{a}_1$ and on $\hat{a}_2$ in this loop are equal to zero, four constraint equations are given as

$$\begin{aligned} g_1 &= n_1 \cos(\theta_n) + n_2 \cos(\theta_r) + n_3 \cos(\theta_f) - n_x \\ &= 0 \\ g_2 &= n_1 \sin(\theta_n) + n_2 \sin(\theta_r) + n_3 \sin(\theta_f) - n_y \\ &= 0 \end{aligned} \quad (1)$$

As $(\bar{N}_a + \bar{N}_b) \cdot \hat{b}_2 = n_4$ in Figure 1, another constraint equation is given as

$$g_3 = n_1 \sin(\theta_n - \theta_f) + n_2 \sin(\theta_r - \theta_f) - n_4 \quad (2)$$

From Equation (1) and (2), the equations of three dependent variables θ_r, n_x, n_y are derived as

$$\theta_r = \theta_f + \sin^{-1} \left[\frac{n_4 - n_1 \sin(\theta_n - \theta_f)}{n_2} \right] \quad (3)$$

$$\begin{aligned} n_x &= n_1 \cos(\theta_n) + n_2 \cos(\theta_r) + n_3 \cos(\theta_f) \\ n_y &= n_1 \sin(\theta_n) + n_2 \sin(\theta_r) + n_3 \sin(\theta_f) \end{aligned} \quad (4)$$

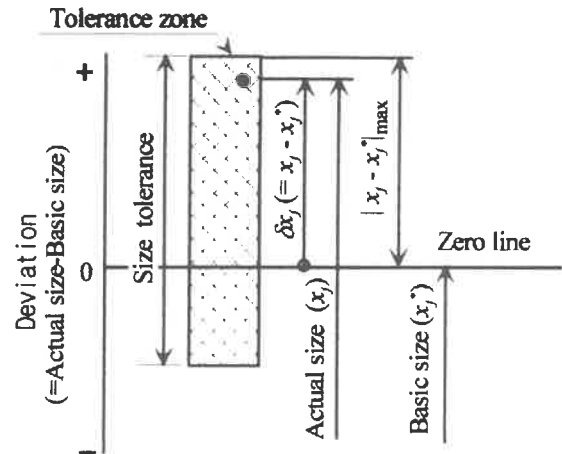


Figure 2. Basic Size, deviation and tolerance in the ISO system (by ISO 286-1:1988(E))

3. Sensitivity analysis for Tolerance

The actual size (x_j) on machine design is never able to be made as the basic size (x_j^*) as shown in Figure 2. Thus, the actual size of a design variable always involves a deviation from the basic size. Also this deviation of each design variable necessarily makes some response errors on dependent variables.

Assuming that dependent variables $\underline{y}: D^m \rightarrow R^n$ have derivatives of all orders everywhere in $\underline{x} \in D^m$, dependent variables $\underline{y}$ can be expanded by Talor expansion near $\underline{y}(\underline{x}^*)$ as

$$y_i(\underline{x}) = y_i(\underline{x}^*) + \sum_{j=1}^m S_{i,j} (x_j - x_j^*) + R_i \quad (5)$$

where $S_{i,j} = \partial y_i / \partial x_j$, R_i is the remainder term.

In Equation (5), if $\underline{x}$ is sufficiently close to $\underline{x}^*$, $R_i \ll \langle y_i(\underline{x}^*) + \sum_{j=1}^m S_{i,j}(\underline{x}^*)(x_j - x_j^*) \rangle$. So, the remainder term can be neglected in Equation (5). Hence, a change in y_i at $\underline{x}^*$ (denoted as δy_i) is rearranged as

$$\begin{aligned} \delta y_i &\approx \sum_{j=1}^m S_{i,j}(\underline{x}^*) \delta x_j \\ &\approx \sum_{j=1}^m S_{i,j}(\underline{x}^*)(x_j - x_j^*) \end{aligned} \quad (6)$$

where δx_j is a small change in x_j^* ($\delta x_j = x_j - x_j^*$), $\delta y_i = y_i(\underline{x}) - y_i(\underline{x}^*)$ and $S_{i,j}(\underline{x}^*)$ is a sensitivity coefficient matrix, which express the sensitivity of i^{th} dependent variable to a small change in j^{th} design variable.

3.1. Maximum response error of dependent variable

When each design variable $\underline{x}$ has a deviation $\delta \underline{x}$, the response error of a dependent variable can be derived from Equation (6) as

$$\begin{aligned} |\delta y_i| &\approx |S_{i,1}(\underline{x}^*)(x_1 - x_1^*) + \dots + S_{i,m}(\underline{x}^*)(x_m - x_m^*)| \\ &\leq |S_{i,1}(\underline{x}^*)||x_1 - x_1^*| + \dots + |S_{i,m}(\underline{x}^*)||x_m - x_m^*| \end{aligned} \quad (7)$$

where $|x_j - x_j^*|$ is the deviation's magnitude of j^{th} design variable.

From Equation (7), it becomes clear that the maximum response error (MRE, $|ME_i|$) of a dependent variable is made when the deviation's magnitude of each design variable is maximum. That is,

$$\begin{aligned} |ME_i| &\approx |S_{i,1}(\underline{x}^*)||x_1 - x_1^*|_{max} + \dots \\ &\quad + |S_{i,m}(\underline{x}^*)||x_m - x_m^*|_{max} \end{aligned} \quad (8)$$

where $|x_j - x_j^*|_{max} (= \max\{|x_j - x_j^*|\})$ is the

deviation's maximum magnitude of j^{th} design variable, and $|ME_i|$ is the MRE of i^{th} dependent variable.

3.2. Procedure for allocation of tolerance

If response errors of dependent variables were successfully derived from maximum deviations of design variables in § 3.1, it might be practicable to reversely allocate tolerances to satisfy with a limit of response error (LRE, $|LE_i|$) of the dependent variable. An allocating procedure of tolerances is introduced here.

First, assuming that every dependent variable has the same maximum deviation in order to simplify a procedure to allocate tolerances, this same maximum deviation is set as an assumed allowable deviation (α) of design variables and can be derived as

$$\begin{aligned} |LE_i| &\approx |S_{i,1}(\underline{x}^*)|\alpha + \dots + |S_{i,m}(\underline{x}^*)|\alpha \\ \therefore \alpha &\approx \frac{|LE_i|}{\sum_{j=1}^m |S_{i,j}(\underline{x}^*)|} \end{aligned} \quad (9)$$

However, as each design variable has its characteristics, it is need to adjust maximum deviation of each design variable to satisfy with its own characteristics. In adjusting process, the maximum response error ($|ME_i|$) is less than or equal to the limit of response error ($|LE_i|$) as

$$|ME_i| \leq |LE_i| \quad (10)$$

Then, the maximum deviation of the h^{th} design variable ($|x_h - x_h^*|_{max}$) can be obtained from changing maximum deviations of other design variables arranging Equation

(9) and (10) as

$$|x_k - x_k^*|_{max} \leq \alpha + \sum_{k=1}^{k-1} \frac{|S_{i,k}(x^*)|}{|S_{i,k}(x^*)|} (\alpha - |x_k - x_k^*|_{max}) \quad (11)$$

$$+ \sum_{l=k+1}^n \frac{|S_{i,l}(x^*)|}{|S_{i,l}(x^*)|} (\alpha - |x_l - x_l^*|_{max})$$

3.3. Sensitivity coefficient matrix of NMM

NMM has four design variables $\underline{\xi} = [n_1, n_2, n_3, n_4]^T$ which are independent each other, and three dependent variables $\underline{g}(\underline{\xi}) = [\theta_f, n_r, n_y]^T$. Here, $\underline{g}$ is the function of $\underline{\xi}$.

There are four constraint equations as shown in Equation (1) and (2) in NMM. These equations can be rearranged as

$$\underline{g}(\underline{\xi}) = 0 \quad (12)$$

where $\underline{g} = [g_1, g_2, g_3]^T$.

Equation (12) can be differentiated with the design variables $\underline{\xi}$ by using the chain rule, and a sensitivity matrix is derived as

$$[\mathbf{r}_{re}] = -[\mathbf{v}_{\alpha\beta}]^{-1}[\mathbf{s}_{\eta\mu}] \quad (13)$$

where $v_{\alpha\beta} = \partial g_\alpha / \partial \xi_\beta$, $r_{re} = \partial r / \partial \xi_e$, $s_{\eta\mu} = \partial g_\mu / \partial \xi_\eta$,

and $[\mathbf{v}_{\alpha\beta}]$ is a nonsingular matrix.

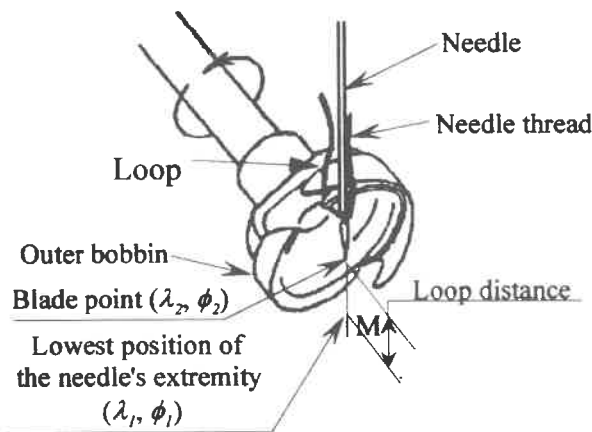


Figure 3. Blade point

A principal function of NMM is to make loops of needle threads. A loop making is influenced by the blade point and the loop distance M as shown in Figure 3. Here, the loop distance M is a distance between the lowest position (λ_1, ϕ_1) of the needle's extremity and the blade point (λ_2, ϕ_2) . Namely,

$$M = \{(\lambda_2 - \lambda_1)^2 + (\phi_2 - \phi_1)^2\}^{1/2} \quad (14)$$

A loop distance M can be differentiated with the design variables $\underline{\xi}$ as

$$\frac{\partial M}{\partial \xi_i} = \frac{(\lambda_2 - \lambda_1)}{M} \left(\frac{\partial \lambda_2}{\partial \xi_i} - \frac{\partial \lambda_1}{\partial \xi_i} \right) + \frac{(\phi_2 - \phi_1)}{M} \left(\frac{\partial \phi_2}{\partial \xi_i} - \frac{\partial \phi_1}{\partial \xi_i} \right) \quad (15)$$

where $\left(\frac{\partial \lambda_1}{\partial \xi_i}, \frac{\partial \phi_1}{\partial \xi_i} \right)$ is a sensitivity of (n_r, n_y) at

the lowest position and $\left(\frac{\partial \lambda_2}{\partial \xi_i}, \frac{\partial \phi_2}{\partial \xi_i} \right)$ is a sensitivity of (n_r, n_y) at the blade point in a locus of the needle's extremity.

4. Applications and Results

4.1. Kinematic analysis of NMM

n_1	n_2	n_3	n_4	θ_f
15.4 mm	48.5 mm	108.4 mm	0.0 mm	270°

Table 1. Nominal sizes of an example model

The nominal sizes of the feeding control mechanism from an example model as tabulated in Table 1.

While the link cam rotates, the value of the needle's extremity (n_r) always is zero, and the locus of the needle's extremity (n_y)

is shown in Figure 4. The needle's extremity is moving downward during $360^\circ \leq \theta_n \leq 270^\circ$, is moving upward during $270^\circ \leq \theta_n \leq 90^\circ$, and is moving downward again during $90^\circ \leq \theta_n \leq 0^\circ$. Thus, the needle's extremity is in the lowest position ($\lambda_1 = 0.0 \text{ mm}$, $\phi_1 = -1723 \text{ mm}$) at $\theta_n = 270^\circ$. In case of the example model, the loop distance M is about 2.2 mm . Thus, the blade point occurs at $\theta_n = 243^\circ$ from the analysis.

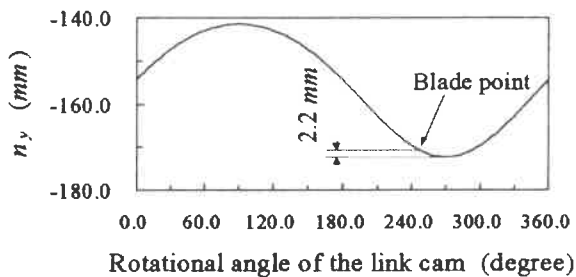


Figure 4. Motion of the needle's extremity(n_y)

4.2. Sensitivity analysis of NMM

ξ_i	n_1	n_2	n_3	n_4
$\partial \lambda_2 / \partial \xi_i \text{ (mm/mm)}$	0.00	0.00	0.00	1.00
$\partial \phi_2 / \partial \xi_i \text{ (mm/mm)}$	-0.82	-1.01	-1.00	0.15

Table 2. Sensitivity coefficients of the blade point (λ_2, ϕ_2)

ξ_i	n_1	n_2	n_3	n_4
$\partial M / \partial \xi_i \text{ (mm/mm)}$	0.18	-0.01	0.00	0.14

Table 3. Sensitivity coefficients of the loop distance M

A sensitivity analysis of the blade point and the loop distance M is investigated here. A sensitivity coefficient matrix (from Equation (13)) at the blade point (λ_2, ϕ_2)

and that (from Equation (15)) at the loop distance M are shown in Table 2 and Table 3.

In order to allocate tolerances of the design variables $\xi = [n_1, n_2, n_3, n_4]^T$, this paper arbitrarily chooses a limit of response errors (LREs) as follows.

CONDITION 1 : The LREs of λ_2 and ϕ_2 on the blade point (λ_2, ϕ_2) are 0.15 mm each other.

CONDITION 2 : The LRE of the loop distance M is 0.11 mm .

From Table 1, the nominal sizes of design variables n_3 is larger than the others. In general case, the bigger size of machine part, the more difficult to get the precise tolerance values. Thus, we try to make larger maximum deviations of n_3 .

An assumed allowable deviation (α) for λ_2 is 0.15 mm , α for ϕ_2 is 0.05 mm , and α for M is 0.33 mm , which are calculated by Equation (9). Because α for ϕ_2 is the smallest value, it is safe to use the sensitivity coefficients of ϕ_2 in order to satisfy with above two CONDITIONS.

In order to make larger maximum deviations of n_3 within LRE of ϕ_2 , maximum deviations of $n_1 (= 0.03 \text{ mm})$, $n_2 (= 0.04 \text{ mm})$ and $n_4 (= 0.04 \text{ mm})$ are arbitrarily set smaller than $\alpha (= 0.05 \text{ mm})$ of ϕ_2 . Thus, maximum deviation of $n_3 (= 0.08 \text{ mm})$ is obtained from Equation (11).

When above determined maximum deviations of design variables are applied to

Equation (4), the worst responses of (n_x, n_y) are represented in Figure 5.

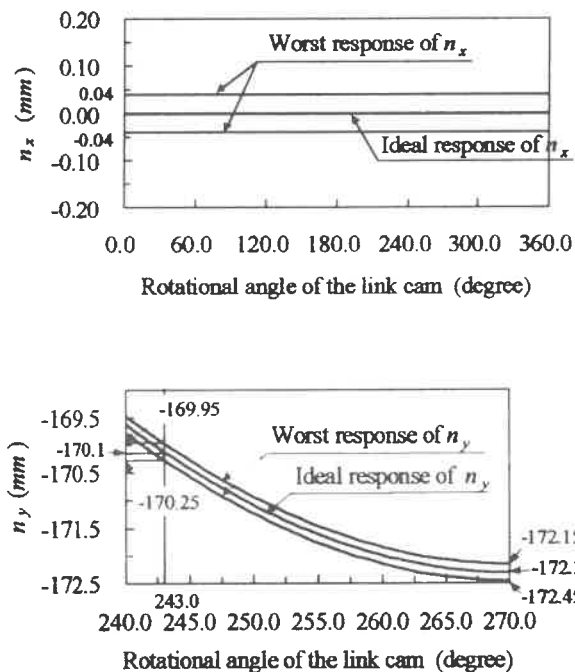


Figure 5. Motion of the needle's extremity when the maximum deviations of design variables are considered

As shown in Figure 5, the maximum deviation of λ_2 is ± 0.04 mm, that of ϕ_2 is ± 0.15 mm, and that of the loop distance $M (= 2.2$ mm) is about zero. Thus, above two CONDITIONS are satisfied.

5. Conclusions

Machine's function related tolerance theory is studied here by using kinematic and sensitivity analysis.

From the kinematics and sensitivity analysis, motions of a machine could be characterized, and maximum deviations of dependent variables might be predicted when the tolerances of design variables are

applied. Inversely, maximum deviations of dependent variables might be determined for tolerance allocation when the a limit of response error of the dependent variable is preset. Therefore, this tolerance theory might be very useful to machine designs, whose functions are much related to kinematics like sewing machines, robots, etc..

References

1. Hartenberg R. S., and Denavit J., Kinematic Synthesis of Linkages, McGraw-Hill, New York, 1964, 315-320
2. L. F. Knappe, A technique for analyzing MECHANISM TOLERANCES, Machine Design, 1963, 25, 155-157
3. R. E. Garrett, and Allen S. Hall, Jr., Effect of Tolerance and Clearance in Linkage Design, Trans. ASME J. Engng Ind., Series B, 1969, 91, 198-202
4. J. Chakraborty, Synthesis of Mechanical Error in Linkage, Mechanism and Machine Theory, 1975, 10, 155-165
5. M. Choubey, and A. C. Rao, Synthesizing Linkages with Minimal Structural and Mechanical Error Based upon Tolerance Allocation, Mechanism and Machine Theory, 1982, 17(2), 91-97
6. Keyoung Jin Chun, Dae Young Shin, "A Study on Feeding Control Mechanism Analysis of Industrial Lock Stitch Sewing Machine", Proc ASPE, 1996, 300-305

A Study on Performance Improvement of Non-Contact Type Seals for High Speed Spindle

Byung Chul Na*, Keyoung Jin Chun, Dong-Chul Han*****

*Korea Automotive Technology Institute, 74 Yongjong-ri Pungse-myon, ChunAnSi
ChungChongnam-do 333-910, KOREA

**Korea Institute of Industrial Technology, San. 17-1, HongChonRi, IbjangMyun, ChonAnSi,
330-820, Korea

***Seoul National University, Seoul 151-742, KOREA

1. Introduction

High speed spindle requires special type of sealing mechanism that prevents a leakage of oil jet or oil mist lubrication system[1]. In the previous study[2], protective collar type labyrinth seal and air jet type seal show more preferable for non-contact types sealing mechanisms working with oil mist or oil jet fluids. Minimum clearance of protective collar type labyrinth seal should be located between inlet side wall and collar, while its size should be minimized taking into account thermal expansions and deflection of the spindle. Therefore this clearance is minimum distance not to be close together physically.

Current work emphasizes on the investigation of the air jet effect on the protective collar type labyrinth seal. To improve the sealing capability of conventional labyrinth seal, air jet is injected against the leakage flow. In this study, both of numerical analysis by CFD(Computational Fluid Dynamics) and experimental measurements are carried out to verify sealing improvements.

Gas or liquid has been used as a working fluid for most of non-contact types seals including the labyrinth seal[3]. However, it is more reasonable to regard two phase flow because oil mist or oil jet are used for high speed spindle's lubrication. In this study, working fluid is regard as two phases that are mixed flow of oil and air phase. Both of turbulence and compressible flow model are also introduced in a CFD analysis to represent

an isentropic process[4,5]. Estimation of non-leaking property is determined by amount of pressure drop in the leakage path. Sealing effect of the leakage clearance and air jet magnitude are studied for various parameters in experiment.

2. Structure of sealing enhancement mechanism

Fig. 1 shows a schematic geometry of an applied seal model. It has a combined geometry of a protective collar type and an air jet type seal. Protective collar that makes a minimum clearance between collar and cavity wall is located at the inlet side of cavity. Then air is injected against leakage flow in the minimum clearance. The isentropic process in the clearance is increased because a clearance is reduced by blocking air jet. It makes enhanced sealing performance along with increasing pressure drop. A clearance reduction effect can be expected in this model even though the clearance can not be close together physically anymore. Fig. 2 shows an effect of clearance reduction by a restrict jet. When the air jet is injected into the minimum clearance, the jetting air reduces an effective leakage clearance. This model can achieve excellent sealing performance even though a rotating spindle and cavity has wide leakage clearance unavoidably. In addition, it can control sealing performance by adjusting the jetting air.

3. Performance testing device

An experimental device is constructed to

test sealing performances as shown in Fig. 3. This device has oil mist generating nozzles to make oil mist lubrication environments. Resolution of pressure measurement is 1 Pa. Leakage clearance is adjustable and capable of close together to 0.01 mm. Sixteen air jet nozzles are installed to the circumference direction for the restrict air jet. Precision pressure sensor(measuring resolution of 1Pa) and regulator are installed at the inlet side of the device to measure and control of leakage pressure. A pressure manipulator supplies high-pressure air to the sixteen nozzles.

4. Testing method and conditions

The pressure drops between inlet and outlet are measured at the speed of 0, 1000, 2000, 3000 rpm to find initial turbulence effects. This test is to find effects of jetting air and spindle rotation speeds without any leakage for the initial turbulence effect. Diameter of collar is 98mm and end clearance is 2mm. Face clearances between collar and wall are adjustable to 0.25mm, 0.5mm, 0.75mm, 1.0mm, 1.25mm and 1.5mm to find clearance effects on leakage flow. For measuring pressure drops, jetting pressures are adjusted to 9.8KPa, 19.6KPa, 29.4KPa, 39.2KPa, and 49.0KPa to find air jetting effects on constant leakage pressure. Initial leakage pressures are ranged from 10Pa to 200Pa to find an initial leakage effect on constant jetting pressure. The estimation of non-leaking property is determined by the amount of pressure drop in the leakage path.

5. Results and discussion

Fig. 4 indicates effect of spindle rotation on constant jetting pressure. This figure is only one case of leakage clearance at 0.25mm while total experiment varies it from 0.25mm to 1.5mm by a clearance adjusting screw. The result shows that rotation speed is not so much affect to pressure drop. Because of a low density of working fluid, turbulence is fully developed at the beginning of test by air

jet and spindle rotation. Therefore increase of circumference direction of rotation speed is not so much affect to pressure drop.

Fig. 5 shows an effect of jetting pressure on constant leakage pressure. In this case, the clearance is 0.5mm. Leakage pressure, represented here, is proportional to the leakage flow. That means a constant leakage flow occurs at constant leakage pressure at the same geometry. The result shows that increase of jetting pressures increase total pressure drops. Thus, sealing performance is improved by the restrict jet. This is very effective way of sealing enhancement in oil mist lubrication system.

Fig. 6 shows an effect of leakage pressure on constant jetting pressure. It shows that a quantity of pressure drop increases between inlet and outlet by restricts jet when it has constant leakage pressure. As shown in dot line of Fig. 6, jetting 39.2 KPa of restrict jet generates 2.1 KPa of total pressure drop where the initial pressure drop(the same as leakage pressure) remains as 0.08 KPa.

Fig. 7 shows an effect of leakage clearance on a constant jetting pressure. If a leakage clearance becomes larger, pressure drops decrease, which representing poor sealing performances in ordinary non-contact type seals. However, this figure shows a inverse behavior that increase of clearance makes bigger pressure drop at clearance of 0.25mm and 0.5mm because of effect of jetting air amount. As a diameter of jetting nozzles are 1.0 mm, leakage clearance of 0.25mm is not enough to send a restrict jet. However, in the case of 0.5mm leakage clearance, effects of restrict jets become dominant to generate pressure drops. Therefore, the restrict jet can achieve good sealing performance although a spindle lubrication system has large leakage clearance inevitably.

Fig. 8 shows an effect of jet nozzle number(4 to 16) on constant leakage pressure. For an ideal case, line source jet is preferable than any number of point source. However, the line source jet is impossible to practical cases. Therefore the effect of jet point number is

investigated as a practical design factor. The effect of jet nozzle number is almost proportional to pressure drop in the range of testing condition. Increase of jet nozzle number means an increase of jet flow, thus, sealing performance is improved until it reaches to line source jet condition.

6. Comparison of CFD Analysis

Boundary conditions are fitted to the testing equipment, and assuming a divided geometry into 16 segment, and each segment has one jetting nozzle. Both of the turbulence and compressible model are introduced in adiabatic condition at 20 °C. The initial condition for comparing with experiments is 0.75mm of leakage clearance and 0.08KPa of leakage pressure where leakage flow rate is 1.3 l/s and corresponding inlet velocity is 0.536m/s. Jet pressures are applied to 9.8KPa, 19.6KPa, 29.4KPa and 39.2KPa as same as the experiment and their flow rates on each nozzle correspond to 0.16l/s, 0.22l/s, 0.27l/s and 0.33l/s respectively. Jet velocity is calculated according to jet flow rate and applied to CFD analysis as jet velocity.

Fig. 9 shows pressure distribution without air jet conditions. This figure show that pressure drop mostly occurs at the entrance of minimum clearance between cavity and collar. Therefore, sealing is achieved by flow that passes through the minimum clearance undergoing an isentropic process. Fig. 10 shows pressure distribution of air jet applied model. This figure show that pressure drop mostly occurs in the minimum clearance near the jet point. It means reductions of leakage clearance by restrict jet are made and flow undergoes more isentropic process. The applied model generates 1.8KPa of pressure drop while no air jet condition generates 0.042KPa of pressure drop. As same results of the experiments, the CFD analysis shows the applied model has excellent sealing

performances compare to protective collar type seals without air jetting.

Fig. 11 shows comparison of experiment and CFD results at the condition of 0.75mm clearance and 0.08KPa leakage pressure. It compares the relative behaviors of pressure drops for the CFD analysis and experiments. Here, CFD results are fitted to experiments by introducing a flow coefficient(0.5). As seen in the Fig. 11, behaviors of pressure drop in experiments and simulation are well match with each other. It verifies an improvement of sealing capability by the restrict jet.

7. Conclusion

This paper offers a way of performance improvements of labyrinth seal used in a high speed machining center. Performance improvements are verified by experiments and the CFD analysis. An applied model has a combined geometry of a protective collar type and air jet type. Effects of sealing improvement are explained as decreasing of leakage clearance by air jetting. Their sealing characteristics are investigated through experiments of a special testing device. Jet flow, which affects sealing characteristics, should be large to achieve an enhanced sealing. Rotation speed does not affect so much to sealing characteristics on this experimental condition. The restrict jet can achieve good sealing performances although a spindle lubrication system which has large leakage clearance inevitably. Increasing jet nozzle number improves sealing characteristics until it reaches to line source condition. The CFD analysis considers turbulence, compressible and two phase flow to simulates a leakage flow of oil mist lubrication system. Behaviors of pressure drop in experiments and simulation are well match with each other. It verifies an improvement of sealing capability by the restrict jet.

References

1. "Seals and Sealing HANDBOOK 2nd Edition", Trade & Technical Press Limited, 1986
2. Byung Chul Na, Keyoung Jin Chun, Dong-Chul Han, "Analysis of Seal Leakage Characteristics for High Speed Spindle", The 6th International Symposium on Phenomena and Dynamics of Rotating Machinery, Vol. 1, pp. 542-551, Feb., 1996
3. Rhode, D. L., Ko, S. H., and Morrison, G. L., "Numerical and Experimental Evaluation of a New Low-Leakage Labyrinth Seal," AIAA Paper 88-2884, July 1988.
4. "PHOENICS Training Course Notes CHAM TR/300", CHAM Limited, 1990
5. Partankar, "Numerical Heat Transfer and Fluid Flow", Spalding, 1980

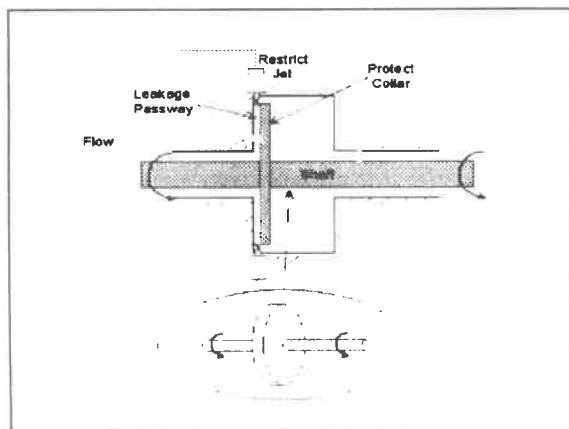


Fig. 1 Applied Seal Model

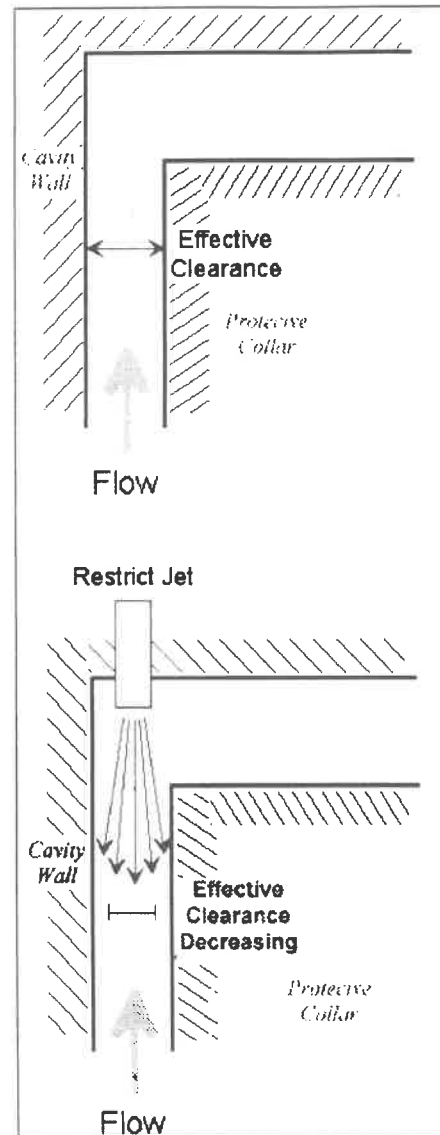


Fig. 2 Effective Clearance Decreasing

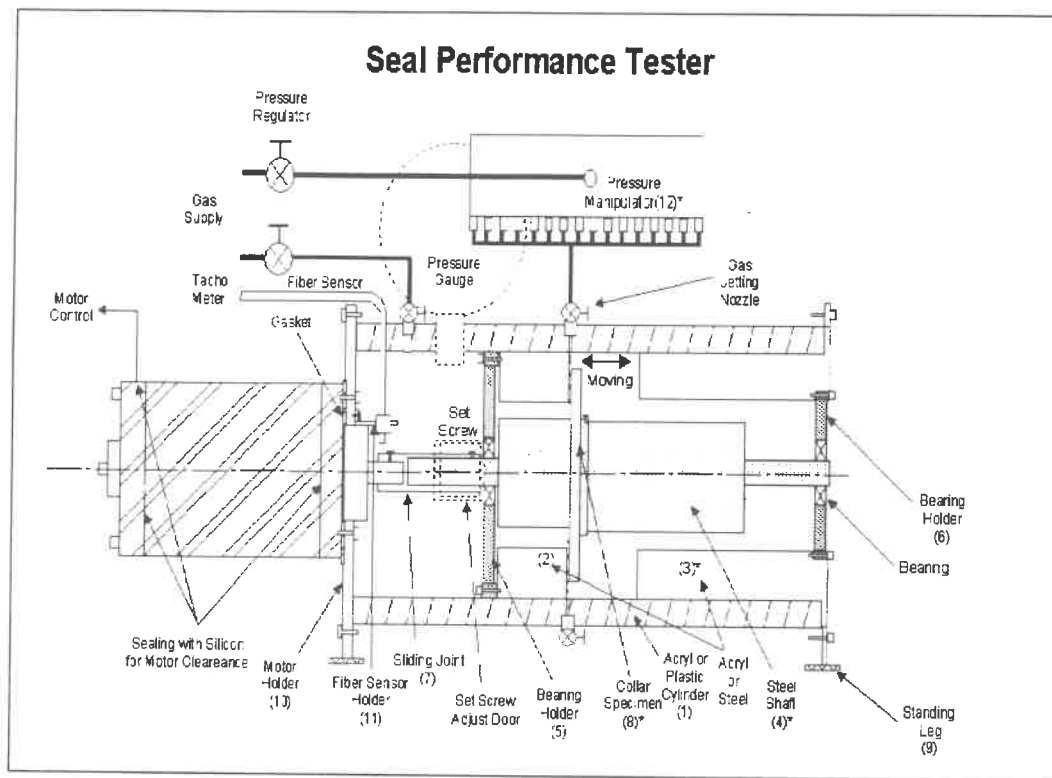


Fig. 3 Performance Testing Device

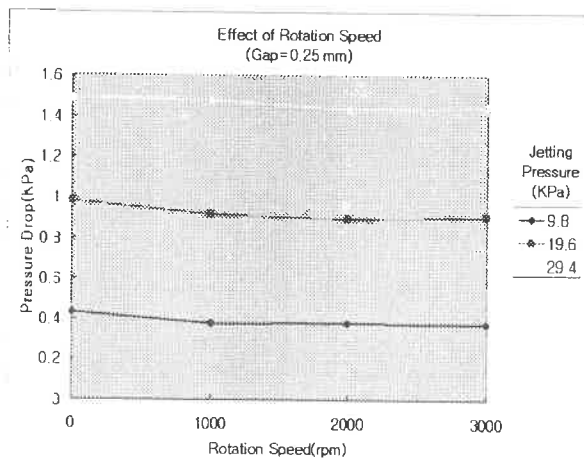


Fig. 4 Effect of Rotation Speed

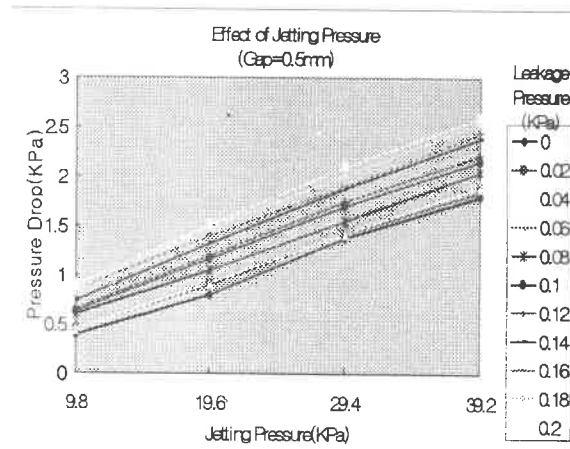


Fig. 5 Effect of Jetting Pressure

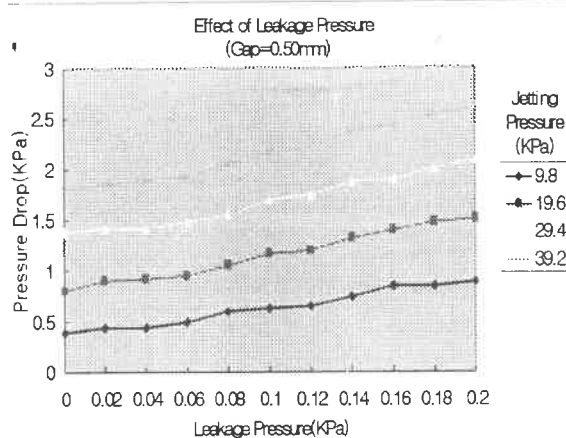


Fig. 6 Effect of leakage Pressure

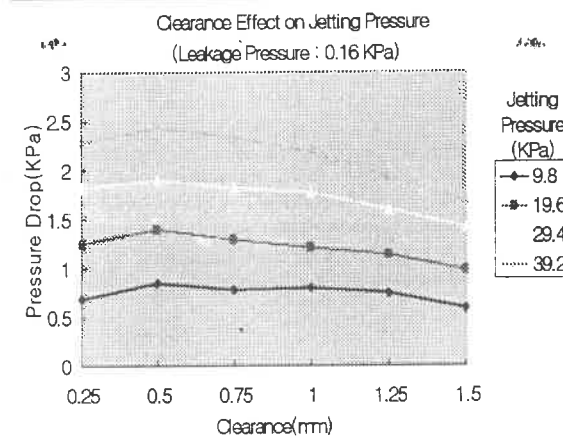


Fig. 7 Effect of Clearance

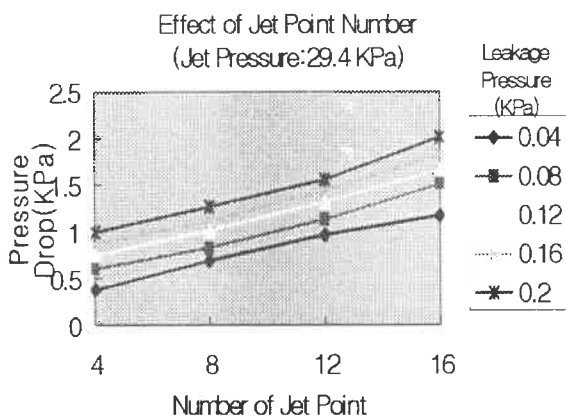


Fig. 8 Effect of Jet Point Number

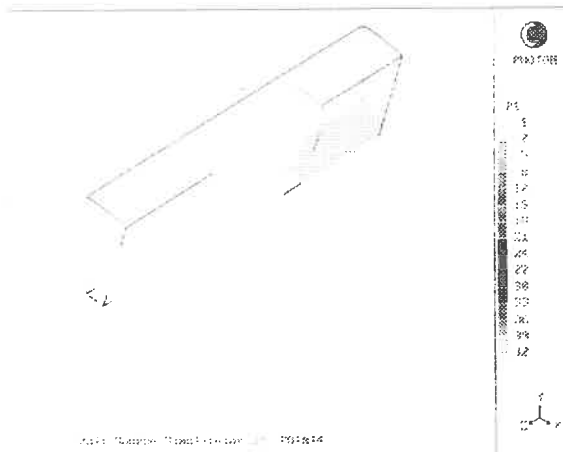


Fig. 9. Pressure Distribution without Air Jet

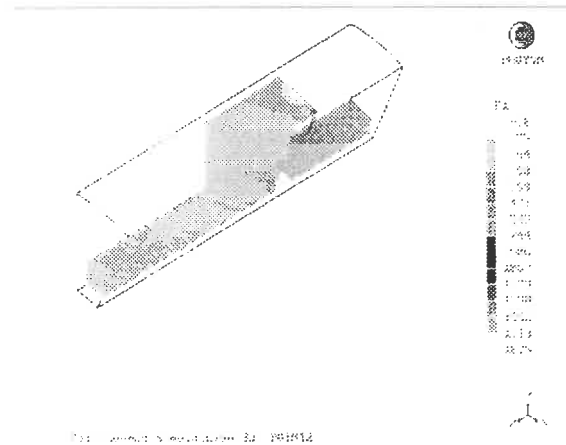


Fig. 10. Pressure Distribution with Air Jet

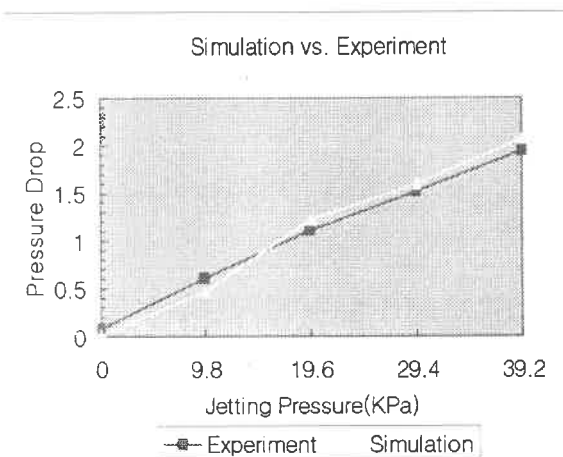


Fig. 11 Pressure Drop of Simulation & Experiment

An End Effector Using Parallel Mechanism for Scanning Electrical Discharge Machining

Katsushi FURUTANI, Kazuhiro SHIBATANI*, Naoki ITOH** and Naotake MOHRI

Toyota Technological Institute

12-1, Hisakata 2-chome, Tempaku-ku, Nagoya 468 JAPAN, e-mail: furutani@toyota-ti.ac.jp

* Tsubakimoto Chain Co. at present, ** Toyoda Machine Works, Ltd. at present

1. Introduction

The electrical discharge machining (EDM) method by scanning an electrode with a primitive shape is proposed to machine a complex shapes in a short time^{1,2)}. Hereafter, this method is called the scanning EDM.

In order to make machining systems flexible, an EDM robot has been proposed³⁾. In this system, an end effector is necessary to obtain high response. Conventional end effectors for EDM have only one degree of freedom (DOF) in the feeding direction⁴⁾. On the other hand, many parallel mechanisms have been studied recently⁵⁾. Since they have multiple DOF, high response and high stiffness, they satisfy the requirements for the end effectors for EDM⁶⁾.

In this paper, an end effector with a parallel mechanism for the scanning EDM is proposed. Then the structure and control method are described. Finally, some results of the scanning EDM are discussed by comparing the motion accuracy of the end effector with its machining accuracy.

2. Parallel mechanism for EDM

2.1 Structure

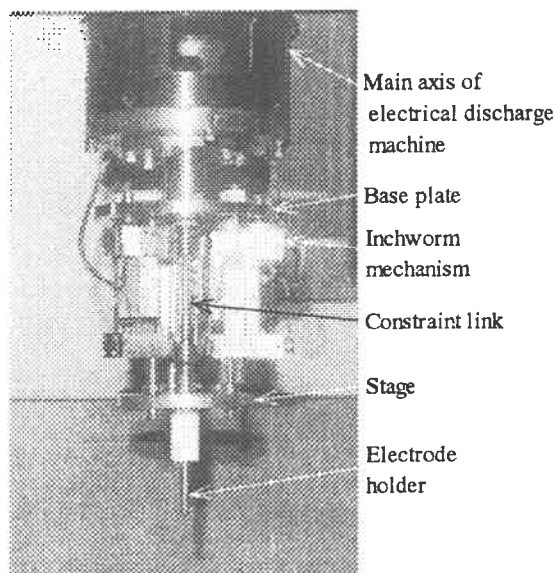


Fig. 1 Appearance of end effector with parallel mechanism

Fig. 1 shows an appearance of a prototype. It consists of a base plate, a stage, three Inchworm⁷⁾ devices performing as links and a constraint link.

This end effector measures $\phi 182 \times 228$ mm. The whole device and the stage weight 6 kg and 2 kg, respectively.

The clamp B and the bottom end of a shaft of each inchworm device are connected with the base plate and the stage with ball joints, respectively. A spline shaft used for the constraint link restricts the motion of the end effector to three DOF. The upper end of the spline shaft is connected with the base plate directly and the lower end, with the stage through a universal joint. Therefore, the end effector can tilt around the x and y axes and it can also move in the z direction.

The Inchworm device measures $120 \times 30 \times 64$ mm and weights 0.8 kg. The shaft of the Inchworm device can move for 30 mm. The piezoelectric elements in the clamps deform $30 \mu\text{m}$ and that in the driver deforms $40 \mu\text{m}$ at the applied voltage of 600 V. In order to support the stage in case of turning off the power, the clamps release the shafts with the diameter of 10 mm

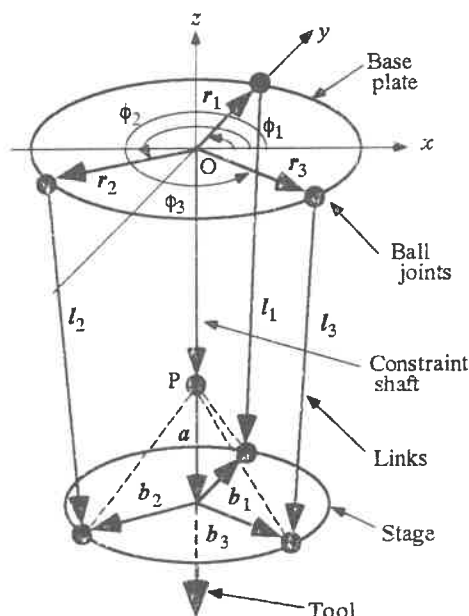


Fig. 2 Geometrical arrangement of links, stage and base plate

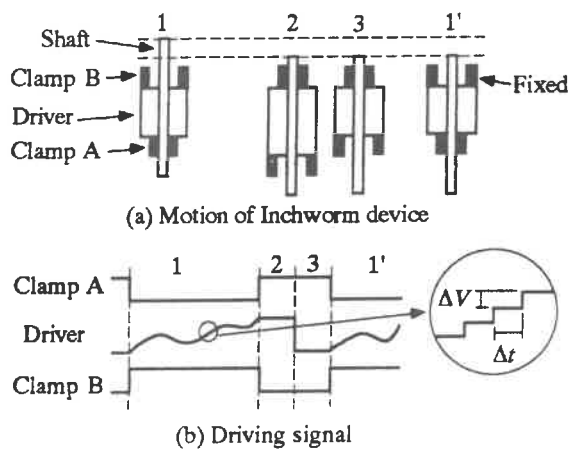


Fig. 3 Driving method by hybrid mode

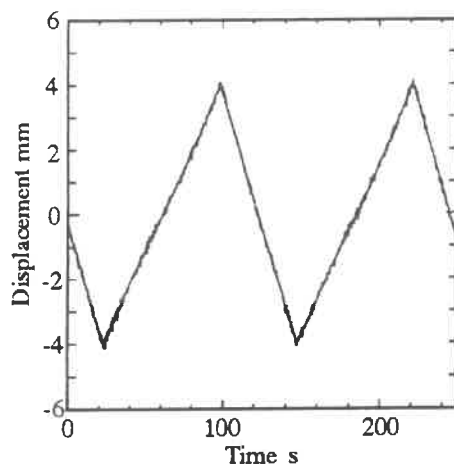


Fig. 4 Coarse motion by step mode

in case of applying the voltage to the piezo. The clamps can hold the shafts up to 30 N.

Potentiometers with the measurement range of 50mm and the straightness of 3% are used to measure the displacement of the shafts of the Inchworm devices.

2.2 Inverse problem

In general, the inverse motion of the Stewart platform type parallel mechanism can be easily calculated by using the geometrical relation. Fig. 2 illustrates a geometrical relation of this end effector. The origin of the coordinates system is the center point of the base plate O. Each solid circle indicates a ball joint with two DOF. The point P indicates a universal joint.

The position and the inclination of the stage are represented by the vector $\mathbf{d}$ and the normal vector of the stage $\mathbf{n} = (n_x, n_y, n_z)$.

The position vector of each link $\mathbf{l}_i$ can be calculated by

$$\mathbf{l}_i = R(\mathbf{a} + \mathbf{b}_i) + \mathbf{d} - \mathbf{r}_i \quad (1)$$

where the vectors $\mathbf{a}$, $\mathbf{b}_i$, $\mathbf{r}_i$ are known.

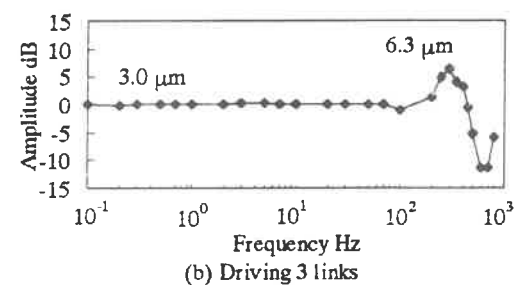
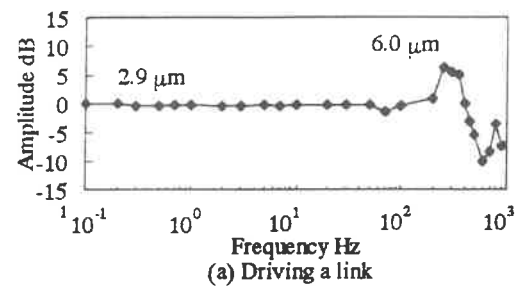


Fig. 5 Frequency response by linear mode

2.3 Driving method

This end effector is driven by the hybrid mode which is a combination of the step mode and the linear mode as shown in Fig. 3. In the step mode, the shaft is driven by the original operation of the Inchworm mechanism. The shafts are positioned roughly. In the linear mode, the piezo for the driver is deformed continuously to displace the stage. The clamps A of all Inchworm devices grasp the shafts and the clamps B of all Inchworm devices release them during the linear mode. The high response can be obtained within a steplike movement in the step mode. The stage is mainly driven by the linear mode during precise positioning or EDM. The shafts are driven by the step mode in coarse motion. When the shaft reaches a reference position, they are driven by the linear mode again.

2.4 Motion performance of end effector

Fig. 4 shows a displacement of the stage by the step mode. The reference position is given every 0.5 mm in the PTP (point to point) control. The stage moves downward faster than upward because of the gravity. The step of the stage is 10 μm in case of the voltage amplitude of the driver of 400 V and the driving frequency of 100 Hz. The positioning accuracy of the stage in the step mode is restricted to ± 30 μm due to the noise from the potentiometers.

Fig. 5 shows the frequency response in which the response curves for one device driving and three devices synchronously driving by linear mode are indicated. The amplitude of 3 μm at 0.1 Hz is indicated as 0dB. In both cases, the resonance

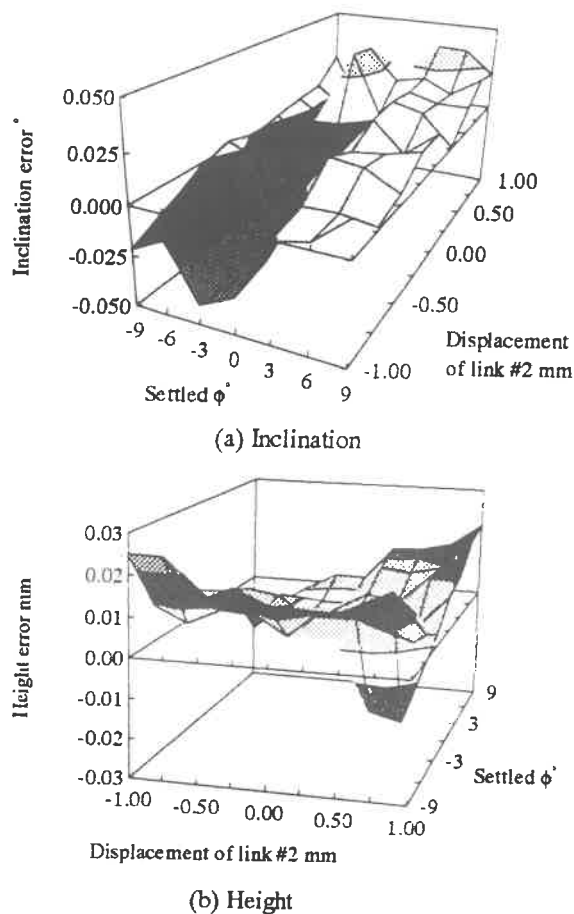


Fig. 6 Positioning error of end effector

frequency is 250 Hz. Therefore, the applicable frequency range is up to 200 Hz as shown in Fig. 5. This end effector driven by the linear mode has enough frequency range for EDM. The clamps grasp the shafts without slip during these experiments.

2.5 Positioning accuracy

Since the displacement range by the linear mode is 10 μ m, the motion accuracy in the step mode affects the machining accuracy more than that in the linear mode. The motion accuracy in the step mode is measured as follows.

Fig. 6 shows the positioning errors in height and inclination which are the difference between the calculated position and the measured one. The instructed inclination of the stage is changed with every 3° from -9 to 9° and the length of one of the links is changed from -1 to 1 mm around the calculated length based on the instructed inclination. In this paper, the length of the links in the forward problem is calculated by the repetitive calculation by Newton method. In all cases, the motion errors are less than $\pm 0.04^\circ$ in inclination and ± 0.03 mm in height along the z direction. The error of the potentiometers of 30 μ m yields the motion errors.

In case of mounting the electrode with the length of 100 mm on the stage, the positioning error of the top of the electrode becomes less than ± 0.09 mm.

3. Scanning EDM

3.1 System configuration

Fig. 7 shows the system configuration for the scanning EDM with the parallel mechanism.

The path of the electrode is controlled by PTP with the position and inclination data of the stage (P/I data). The driving signals are supplied through amplifiers for the piezos. The electrode is controlled according to the forward/backward signal (F/B signal) decided by comparing the average supply currents with a certain reference current. The electrode moves to the next reference position in case of the forward signal and reverse to the previous reference position in case of the

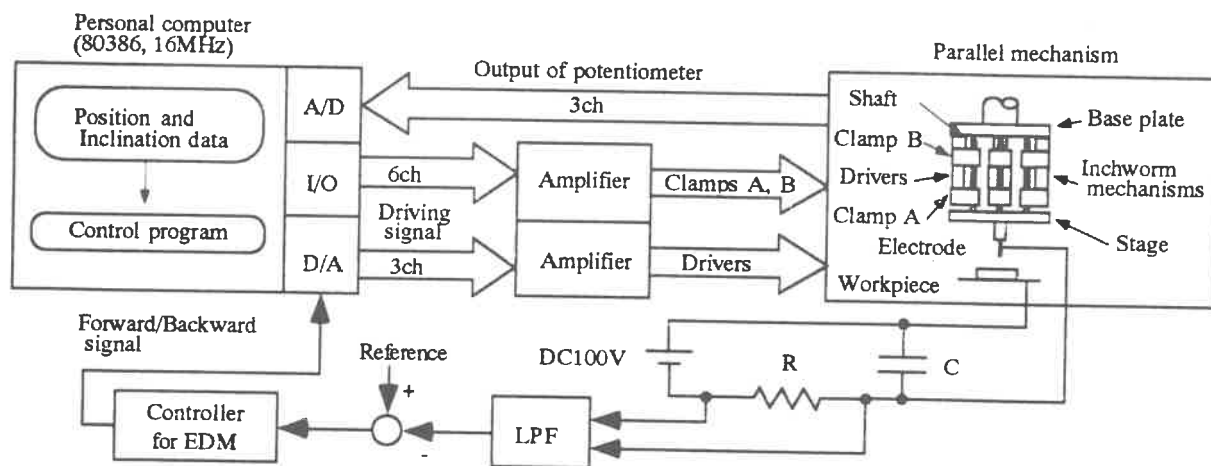


Fig. 7 System configuration of scanning EDM with parallel mechanism

backward signal in order to reduce the machining error and to avoid a short-circuited state. The displacement direction of the shaft is decided according to the current position and F/B signal. The F/B signal has the priority to the P/I data for stable EDM.

3.2 Machining grooves

In order to compare the machining accuracy with the motion accuracy, grooves are machined on AISI 304 stainless steel plates.

The discharge circuit used in the experiments consists of the capacitor of 1 μ F and the resistance 168 Ω . The gap voltage is settled to 100 V. A tungsten rod of 1.0mm in diameter is used for an electrode. The driving frequencies are settled to 25 Hz for the step mode and 150 Hz for the linear mode.

The reference points are given at every 0.1 mm for the straight grooves and at every 2° for the circular groove. After the electrode is scanned along the whole part of the groove, the stage is driven by 0.1 mm in the z direction.

Fig. 8 (a) shows a result of machined crosslike groove of 5mm in length. The machining time is approximately 30 minutes. The width and the average depth of the grooves are 1.10 mm and 0.58 mm, respectively. Though the settled angle between two grooves is 90.0°, the actual angle is 90.8°.

Fig. 8 (b) shows a result of machined circular groove when the center line of the electrode draws a circle of 5 mm in diameter. The machining time is approximately 70 minutes. The outer and the inner radii are 3.210 mm and 2.015 mm, respectively. The average depth of the groove is 0.31 mm. The roundness is 0.142 mm for the outer circle and 0.126 mm for the inner circle. The distance between the centers of the outer and the inner circles is 0.043 mm. Since the machining accuracy by conventional EDM is approximately several micrometers and the machining accuracy is less than the motion accuracy measured in Section 2, the machining accuracy depends on the motion accuracy.

4. Conclusions

The end effector with three DOF by using a parallel mechanism is developed and applied to EDM.

The following conclusion can be drawn through the experiments:

- (1) The accuracy of machined grooves is less than the motion accuracy. The surface modification with silicon is demonstrated.
- (2) This end effector has an enough frequency range for EDM. The frequency range in linear mode is up to 200Hz.

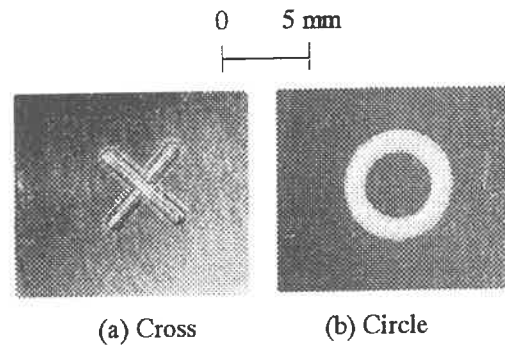


Fig. 8 Examples of machined shapes

Acknowledgment

The authors wish to thank to the Ceramic R&D Dept. of Denso Corp. for providing the piezoelectric elements and Mr. N. Andoh who graduated from Toyota Technological Institute for their helpful cooperation. This study is financially supported by the Die and Mold Technology Promotion Foundation, Amada Foundation for Machine Tool Technology, Tokai Foundation for Technology, OSG Fund and Grant-in-aid for Scientific Research (A) (1) (06555041) and (C) (2) (07650155) of the Ministry of Education, Science, Sports and Culture, Japan.

References

- 1) T. Kaneko et al.: Improvement of 3D NC Contouring EDM Using Cylindrical Electrode, Proc. Int. Sympo. Electro Machining 9, Nagoya, Japan, pp. 49-52 (1989).
- 2) K. Furutani et al.: 3D Profile Generation by Dot-Matrix Method Using Miniaturized Electrode Feeding Devices, Proc. 11th Ann. Meeting of Am. Soc. Prec. Eng., pp. 86-90 (1996).
- 3) T. Narikiyo et al.: Control System Design for Macro/ Micro Manipulator with Application to Electro-discharge Machining, Proc. IEEE/ RSJ/ GI Int. Conf. Intelligent Rob. Syst., 2, pp. 1454-1460 (1994).
- 4) D. Kremer et al.: Effects of Ultrasonic Vibrations on the Performances in EDM, Ann. CIRP, 38, 1, pp. 199-202 (1989).
- 5) R. S. Stoughton et al.: A Modified Stewart Platform Manipulator with Improved Dexterity, IEEE Trans. Rob. Autom., 9, 2, pp. 166-173 (1993).
- 6) K. Furutani et al.: Miniaturized Electrode Feeding Devices Using Piezoelectric Elements, Proc. Int. Symp. Electro-machining, pp. 621-628 (1995).
- 7) Burleigh Instrum. Inc.: Inchworm catalog.

SELF-CALIBRATION FOR A HEXAPOD COORDINATE MEASURING MACHINE

Chin-Chou Huang, John C. Ziegert
Machine Tool Research Center, University of Florida, Gainesville, FL

1. Introduction

Typically a conventional coordinate measuring machine (CMM) uses serial kinematic linkages, each providing a single degree of freedom. The configuration may be a bridge type or cantilever construction. In order to achieve high accuracy measurements, the conventional CMM must be placed in a temperature-controlled environment to minimize the structural distortion due to thermal gradients in the machine elements. Therefore, the total cost for the machine itself and the creation and maintenance of the required working environment is a significant economic issue for CMM users.

An alternative is the hexapod coordinate measuring machine (HCMM) as shown in Figure 1. This machine uses a parallel structure in which light weight telescoping struts are used to reduce the moving masses typically in the conventional CMMs. The HCMM is designed to self calibrate and compensate for thermally induced changes in its own elements. Thus the requirements for temperature control in the operating environment are greatly reduced.

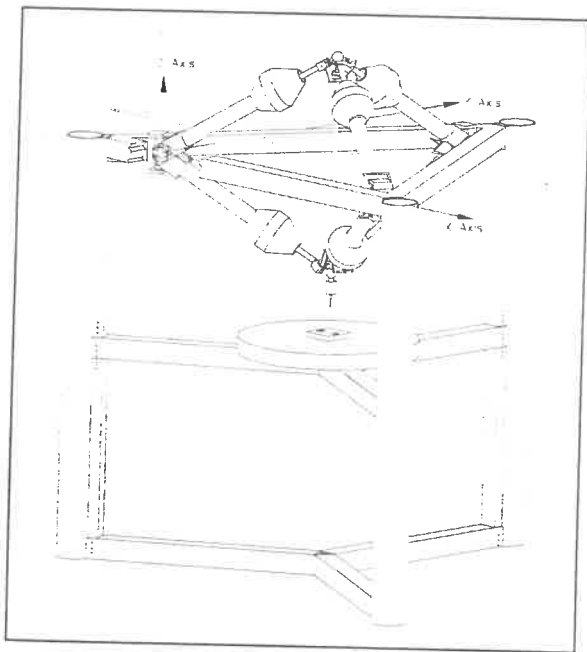


Figure 1 The hexapod coordinate measuring machine

This five degree of freedom HCMM we are now constructing is composed of six telescoping struts with laser interferometric displacement measurement and a single fixed-length center rod. The six telescoping struts are mounted in pairs, with one pair at each of three spherical joints around the base of the machine. The configuration of the machine resembles two tetrahedrons which share a common base and whose apex points are connected by the fixed rod. The laser interferometers measure the length changes of each strut. However, in order to calculate the coordinates of the probe, it is necessary to know the absolute length of each strut and the distance between the three base joints. Thus there are nine kinematic parameters to be determined in the calibration process, the three base lengths and the initial lengths of the six struts.

This paper presents a method to find the initial strut lengths and the base joint locations without adopting any external artifacts or measurements. During calibration the measuring probe is moved to a number of different positions within the measuring volume. For each position, the center rod length is assumed to remain constant. The length is calculated based on an initial guess of the machine kinematic parameters. A non-linear least squares problem is then formulated to search for the values of the base locations and initial strut lengths which minimize the variation in the computed center rod length. The entire calibration procedure can be completed within several seconds after the final movement of probe. Therefore, it is possible to correct in real time the thermally induced changes in the strut lengths and the base positions.

2. HCMM performance

The maximum and minimum strut lengths are 762 and 1270 mm respectively. The expected HCMM system performance was outlined by Jokiel [1] as follows:

1. Overall accuracy: ± 2.2 to $\pm 5.3 \mu\text{m}$.
2. Workspace: A cylindrical volume 381 mm in diameter and 381 mm in height.
3. Speed: Maximum constant probe speed of 762 mm/sec. Strut extension speed of 635 mm/sec.

3. Kinematic Modeling

The origin of the world coordinate system is placed on one of the base sphere centers, BS1. The x axis lies along a line connecting two base sphere centers. The plane passing through the three base spheres is chosen as the x-y plane. The BS2 has coordinate of (r,0,0) and BS3 is located at (b,h,0).

Because of the telescoping strut design, each of the six struts may change its length to satisfy the kinematic constraint when the center rod is moved from one position to the other. Let's say the k th strut, with its initial (or starting) length of $a(k)$, has a length change of $S(k,i)$ at the i th configuration. At this instant its absolute length is expressed as $l(k,i) = a(k) + S(k,i)$ where $k=1,\dots,6$ and $i=1,\dots,m$. The initial length $a(k)$ is a constant value and the length change $S(k,i)$ will vary from configuration to configuration. The centers of the upper and lower spheres, P_u and P_l can be represented as:

$$P_u(i) = f(a_1, a_3, a_5, r, b, h, S(k,i)) \quad P_l(i) = f(a_2, a_4, a_6, r, b, h, S(k,i))$$

Therefore, the center rod length can be calculated as:

$$Lc = \sqrt{(P_{u,x} - P_{l,x})^2 + (P_{u,y} - P_{l,y})^2 + (P_{u,z} - P_{l,z})^2} = f(a_1, \dots, a_6, r, b, h, S(k,i))$$

4. Measurement

For the calibration of a conventional CMM, the measurement process is time consuming, laborious, and prone to human error. In contrast, the measurement process for the HCMM is simply movement of the center rod to a number of different positions within its work volume, and recording the displacement reading in each strut. Therefore, this calibration procedure is called "self-calibration" since no external measurement device is required. Data collection is quick and easy. However, it is necessary to minimize the number of measurement, because we assume that there is no thermal effect during the calibration process. If too many measurements are used, could the time required may allow significant thermal deformation in the structure.

We will show that the self-calibration process is very sensitive to the level of measurement noise. The measurement noise may include the modeling imperfection or measurement error from the laser interferometer. To eliminate or minimize the effect owing to measurement noise, then, becomes one of the most difficult problems in the parameter identification. In general, the magnitude of the noise itself is quite small, but its effect on the estimation accuracy may be rather significant. To minimize the effects of the inevitable noise on the results of estimation, the selection of measurement configurations becomes important

5. Identification

Because of the incorrect nominal values, measurement noise and modeling error, the calculated center rod length from the model will not have the same value at each pose. Therefore, an algorithm called *reference residual* is used in the HCMM calibration procedure. It is formulated as follows: $e_i = Lc_r^2 - \hat{Lc}_i^2$ where Lc_r is an assumed value for the center rod value which is an approximate guess. It is not necessary for this value to be known exactly and it will be treated as the tenth kinematic parameter which needs to be identified. The objective is then to find a set of parameters including the Lc_r to minimize the reference residual, e_i . The least-squares method can be expressed as

$$\text{Minimize } \sum_{i=1}^m e_i^2 \quad \text{Equation 1}$$

Since Equation 1 is highly non-linear, one method to solve it is to linearize this equation. At the i th configuration of the HCMM, by using the Taylor expansion, the reference residual can be written as:

$$e_i = e_i|_{A_0} + \frac{\partial e_i}{\partial a_1} \Delta a_1 + \dots + \frac{\partial e_i}{\partial r} \Delta r + \frac{\partial e_i}{\partial b} \Delta b + \frac{\partial e_i}{\partial h} \Delta h + \frac{\partial e_i}{\partial Lc_r} \Delta Lc_r + \text{higher order terms}$$

where vector A_0 is the nominal value for parameters $\{a_1, \dots, a_6, r, b, h, Lc_r\}^T$.

Suppose that the center rod is placed into m poses, this provides a vector of E which is written as

$$E = E_0 + J\Delta A + \text{higher order terms} \quad \text{Equation 2}$$

where J is referred as Jacobian matrix. If the nominal values of A_0 are close enough to the true values, the higher order terms in Equation 2 can be ignored. In this case the objective function of Equation 1 can be rewritten as:

$$\text{Minimize } f = (E_0 + J\Delta A)^T (E_0 + J\Delta A) \quad \text{Equation 3}$$

Using the iterative linear least-squares method, the minimum value of f can be obtained

$$\Delta A = (J^T J)^{-1} J^T E_o$$

During each iteration the nominal parameter values are updated as

$$A_{new} = A_{old} + \Delta A$$

Equation 4

This procedure is iterated until the largest value in vector ΔA is less than the desired tolerance.

6. Calibration Simulation

A computer simulation of the self calibration process has been implemented. The flowchart for the calibration simulations is shown in Figure 2. The approach taken during the course of the identification simulations

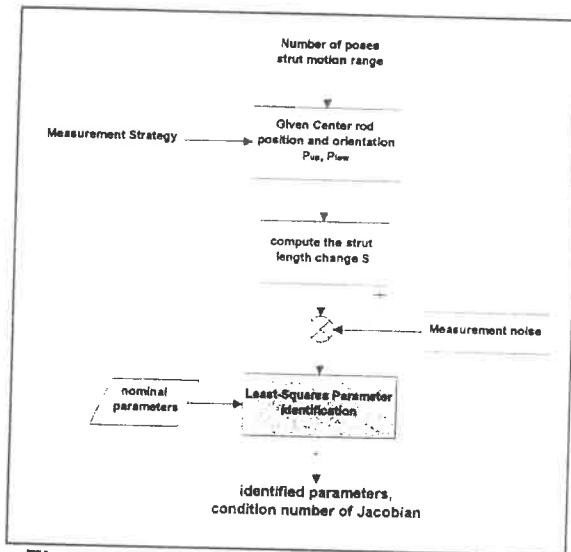


Figure 2 The calibration simulation flowchart

Table 1 The true kinematic parameters (mm)

a1	a2	a3	a4	a5	a6
1016	1016	1016	1016	1016	1016
r	b	h	Lc		
1600	800	800√3	856		

ten nominal parameters in the HCMM model was " δA ". The measurement strategy types are "Strg". After the calibration procedure is completed, the identified parameter errors $\delta a_1, \dots, \delta a_6, \delta r, \delta b, \delta h$, and δLc are represented as " δA_{rms} " which is found by taking the root mean square of all ten parameter errors. The magnitude of lower sphere position error at each pose is " $|\delta P|$ ". The condition number of identification Jacobian is " $Cond(J)$ " which is the value of the maximum singular value of the Jacobian matrix divided by the minimum singular value of the Jacobian matrix.

In our iterative least-squares algorithm each entity in the Jacobian matrix was calculated analytically by taking the partial derivative of the objective function. The iterative procedure was terminated when the maximum value in parameter update vector, ΔA , became less than 2.54×10^{-14} m.

6.1 Measurement Strategy

Five measurement strategies have been examined through computer simulation, where the center rod motion locus is shown in Figure 3. The calibration results are shown in Table 2.

Table 2 Comparison of measurement strategies

	Strg 1	Strg 2	Strg 3	Strg 4	Strg 5
δA_{rms}	not identifiable	not identifiable	60.4 μ m	53.3 μ m	1.7 μ m
cond(J)	ill conditioned	ill conditioned	11540	12917	303
m=200 noise: $\pm 0.254 \mu$ m initial $\delta A = 12.7$ mm					

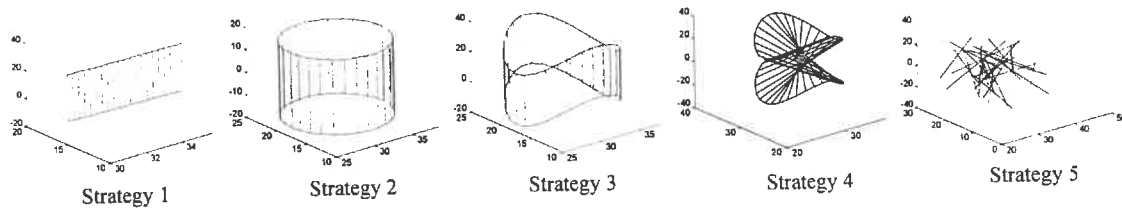


Figure 3 Measurement Strategy

6.2 Pose Selection Each pose set has a sensitivity level for the measurement noise so that using a different pose set would yield different calibration results. To present the noise sensitivity of a pose set, one widely used index is the condition number of the Jacobian matrix J . Figure 4 shows the calibration results of a thousand simulations. One can observe from the simulation results that a pose set with smaller condition number does not guarantee that the calibration results will be better than for a pose set with a higher condition number. However, comparing the upper boundaries of condition numbers one can see the trend toward pose sets with smaller condition numbers having smaller a upper bound in position errors. If one draws a straight line above all the points in Figure 4, one can see the trend. From the above simulation we found that the position error varied from a maximum of $33\mu\text{m}$ ($\text{cond}=1100$) to a minimum of $0.337\mu\text{m}$ ($\text{cond}=393$). However, for a randomly selected pose set the calibration result has a large possibility of being located in the region of the mean value from 1000 simulation results. Thus we choose to use the mean value and its standard deviation of the 1000 simulation results to represent the overall calibration results.

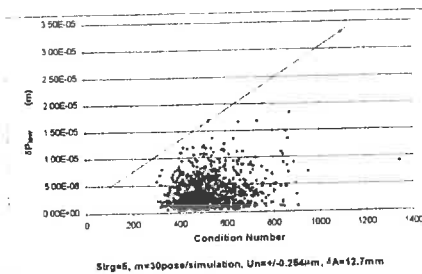


Figure 4 Calibration Results of 1000 Simulations

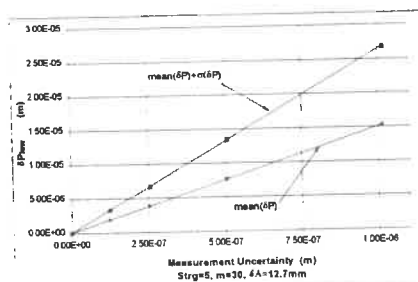


Figure 5 HCM Accuracy vs. Measurement Noise

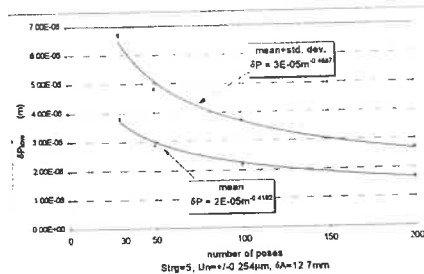


Figure 6 HCM Accuracy vs. Number of Poses

6.3 Effect of Measurement Uncertainty This simulation was repeated using different levels of measurement uncertainty. The results are shown in Figure 5.

6.4 Effect of Number of Poses Figure 6 shows the calibration results obtained using an increasing number of poses.

7. Conclusion

A good measurement strategy will determine the quality of the parametric identification results. In some cases, the parameters may be unidentifiable from a given pose set such as the straight line or circular motion. The best measurement strategy of those tested is random selection. Furthermore, it has been found that not all of the pose sets selected from the random strategy obtained acceptable calibration results. A good calibration result should be obtained when the pose set is least sensitive to measurement noise. It has been shown that increasing the number of poses and reducing the measurement error are two methods to improve the calibration quality. The most significant factor that will improve the calibration quality is the pose set selection. Thus the condition number of the identification matrix can be used as an important indicator, but not an absolute reference, during the selection of the poses. The condition number computed from the experiment can be used to refer back to the simulation tables and estimate the quality of the pose set, and suggest a level of confidence in parameter identification.

References

1. Jokiel, B., "Analysis of a Parallel Architecture Coordinate Measuring Machine," Master's Thesis, University of Florida, 1996.

Three Parameter Solution of Time-varying Thermal Errors

T Gim^{*,**}, J Corbett^{**} and A E Gee^{**}

* Machine Tool R&D Center, Daewoo Heavy Industries, Ltd.
601-3, Namsan-dong, Changwon, Kyungnam, 641-090, Korea (on leave at ** 1994-97)

** School of Industrial & Manufacturing Science
Cranfield University, Cranfield, Bedford, MK43 0AL, UK

Introduction

Errors in machine tools are generally classified as being time or spatially dependent. Thermal errors are strongly dependent on the continuously changing operating conditions of a machine and its surrounding environment, which account for as much as 70% of workpiece inaccuracy, according to Bryan [1]. Uniform temperature rises or stable temperature gradients, which produce time-invariant thermal errors, are considered to be rare in ordinary shop floor environments. Time-varying (dynamic) thermal errors in machine tools, as the result of the effects of continuous temperature variations, had not been investigated satisfactorily, so compensation techniques using empirical correlations had been the main research subject [1,2]. In respect of applications-related design considerations, it was decided that a prime objective of this study should be the development of a simple, yet systematic, physically meaningful modeling method of time-varying thermal errors using appropriate thermal parameters.

Modeling Thermal Errors by Three Parameters

It was considered that accurate analysis of thermal deformation should accommodate the distributed nature of physical and material properties, and afterwards a lumped model might be employed as viable approximation. The distributed nature makes problems difficult because governing equations are complex. From a metrological point of view, we sample frequently the behavior of an object at discrete points considered to be interesting, so that the overall behavior can be described sufficiently with less effort and resource. This is appropriate for the case of thermal error problems of machine tools and coordinate-measuring machines. Errors occurring in machine movement are often measured at discrete points along the path of the movements. Thus, a point-wise description of thermal errors is reasonable in a metrological sense and difficulties in analyzing a distributed system will be overcome. Overall behavior can be obtained by interpolation between points of interest.

In Figure 1, there are several points on the body, denoted by P_i , which are assumed to be points of interest for this particular body. Points P_1 to P_5 are located on the surface of the body, P_6 is between the surface and its center, and P_7 is at its center. Each point has its own response under thermal loadings. This response at a particular time can be accurately described as a block diagram shown in Figure 2. From the theoretical analysis of simple machine elements such as bars, beams and cylinders, and extensive finite-element simulation data for a straightedge subject to room temperature variations [3], three thermal parameters, i.e. time-delay, time-constant and

gain, were identified to obtain a precise description of the thermal deformation of a point of a machine body. A heat input will cause a movement of point P_i to another position, and it has three components in a Cartesian coordinate system. During the movement, four types of process-elements, including processes due to those three thermal parameters, are acting, i.e. a delay element, first-order element, gain element and positive feedback element. This block diagram is the result of analyses done so far.

g : gravitational acceleration
 F_1, F_2 : external forces
 T : temperature
 q_b : heat flux at boundary
 q_r : radiation at boundary
 q_c : convection at boundary
 n : outward normal coordinate
 T_b : known temperature at boundary

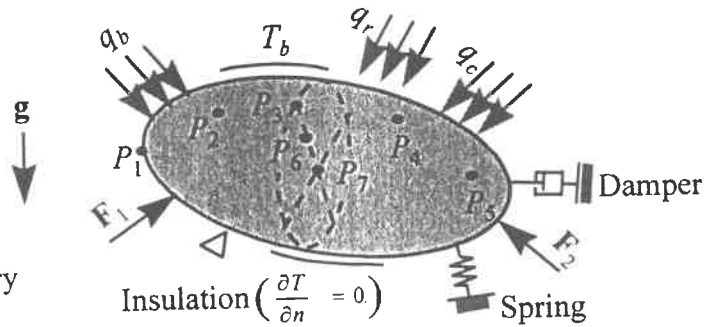


Figure 1 Body subject to external forces and thermal loadings

By definition, a delay or dead-time element causes the effect of an output to be delayed a finite time. In heat-transfer processes there is a transportation delay of heat flux. Point P_6 at the center will experience a greater time delay than points on the surface. This effect is particularly significant for a material having a low thermal diffusivity such as granite. A body having small dimensions and a moderate thermal diffusivity will not exhibit a significant time delay, and the effect of a delay element can be ignored. Thermal environments also affect the magnitude of time delay. Time delay, ranging from fractions of seconds to hours, depends on the location of the point of interest with respect to heat sources and material properties. Thermal environments generally vary with time, and thus time delay varies with time too.

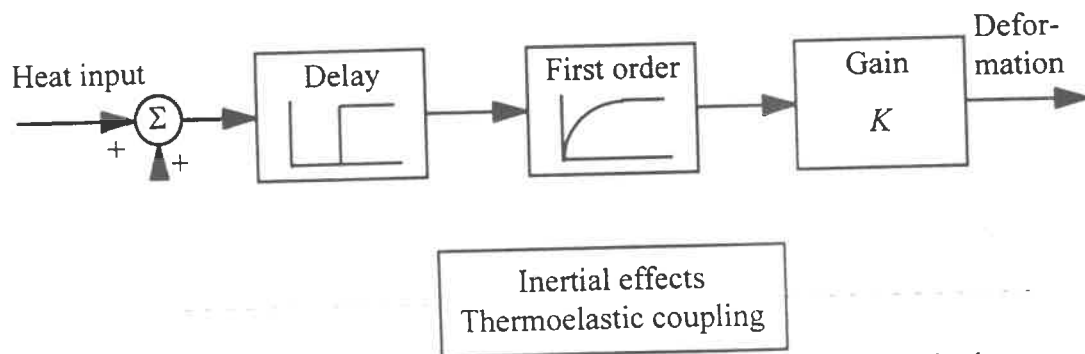


Figure 2 Block diagram of thermal deformation at a point in a body

A first-order element is well recognized in terms of thermal deformations, and characterized by its time-constant. A particular characteristic of a first-order element thermal deformation process is a very large time-constant in comparison with electrical or servo systems, i.e. of the order of hours to several days. Correspondingly, steady-state responses are not of great use for the problems of thermal deformations. It is understood that several factors are acting simultaneously to form the time-constant, i.e. the heat capacity of a body, its size and the thermal environment.

The thermal environment generally varies with time, and thus the time-constant varies with time too.

A 'gain element' is defined as converting changes in temperature to corresponding displacements in which coefficient of thermal expansion plays a principal role in this process. A body with a high thermal expansion coefficient will generally exhibit a high gain. Another important factor in this process is the constraint conditions of the body and they affect the distribution of the gain of each point significantly. A freely-expanding body will have a uniform distribution of the gain both throughout its surface and inside, whereas a constrained one will exhibit significant changes in gain. Finally, thermal environments also affect gain. The thermal environment generally varies with time, and thus gain varies with time too.

As a further element, there is the effect of positive feedback related to two physical phenomena, i.e. inertial effects and thermo-elastic coupling. Inertial effects are associated with the rate of change of strain rate. If an engineering component is stationary (or in a state of uniform motion) and external forces are applied gradually, then inertial effects are negligible. This applies to the case of machine tool components. Thermo-elastic coupling is related to strain rate. A process involving high strain rate, such as tensile testing, produces heat even though there are no heat inputs, but machine tool components do not generally involve such high strain rates. In elastic bodies undergoing non-inertial loading, it is observed that strain rate is always small compared with the rate of change of temperature, and thus we can discard the positive feedback element for this problem. In conclusion, thermal deformation of a body can be understood accurately with a point-wise delay-plus-first-order process, and thus thermal errors can be described by several delay-plus-first-order processes of points of interest in a machine structure.

Refinement of Point-wise Description Method

As a first step, a governing equation of thermal deformation of a point at a particular time will be obtained., then the transfer function G of thermal deformation process is obtained using Figure 2:

$$G = \frac{Ke^{-t_D s}}{1 + \tau s} \quad (1)$$

where K is the gain of the gain element; t_D is the dead time of the time-delay element, t is the time-constant of the first-order element and s is the Laplace variable. This transfer function is equivalent to the following first-order differential equation:

$$\tau \frac{du}{dt} + u = KF_{th}(t - t_D)H(t - t_D) \quad (2)$$

where u is the displacement of a point in a body caused by temperature rise; $F_{th}(t)$ denotes a time-varying thermal loading; $H(t)$ is the Heaviside unit step function. Because system parameters are varying with time due to continuously changing thermal environments and some finite number of sampling points will be involved, the governing equations of the thermal

deformation of a body can be effectively analyzed by employing a state-space analysis method. One benefit of using Equation 2 is that it provides flexibility in choosing system input when analyzing the thermal behavior of a machine, because it describes the input-output characteristics. If the heat input to a machine component is chosen to be the spindle speed, then the gain has a unit of the form [(thermal deflection)(spindle speed)-1].

Evaluation of Thermal Parameters

Completion of this approach was made by presenting the methods by which the three thermal parameters can be evaluated. The first method employs analytical tools based on simplifying assumptions about the shape and boundary conditions of machine components. The second method was to apply numerical techniques to complex machine components. Lastly, the three thermal parameters can be easily identified using popular parameter identification techniques which can be applied to time-varying cases by their recursive forms.

Direct experimental evaluation of the thermal parameters using a full-scale model was investigated using a single-point diamond turning research machine at Cranfield university. The experimental test of thermal errors and parameter evaluation using the test data was performed on the aero-static spindle unit because the main function of the machine is the facing of thin disks. Before the test, the expected thermal deformation of the spindle unit was calculated analytically. In order to identify time-varying system parameters quantitatively, some modified recursive least-squares methods were employed. They were, however, found to be sensitive to noise in the measurement data. The experiment implied that, under constant thermal and operating environment, the time-varying characteristics of the parameters can be neglected because they gave nearly equal thermal parameters under the same heat input conditions.

A scaled physical test model method was investigated regarding its use as a design test method. Also, it can be regarded as an indirect evaluation method of the three thermal parameters employing smaller scaled models as a feasible solution when dealing with large machine tool structures. It was found that there were some problems in directly applying similarity rules in thermal processes. These are due to the very large increases in the heat transfer coefficient and much smaller time scale for test models. This was overcome by using the inverse proportional relationship between the time-constant and heat transfer coefficient.

References

1. Bryan, J.B. (1990), "International status of thermal error research", Ann. CIRP, Vol. 39, No. 2, pp. 645-656
2. Weck, M., McKeown, P., Bonse, R. and Herbst, U. (1995), "Reduction and compensation of thermal errors in machine tools", Ann. CIRP, Vol. 44, No. 2, pp. 589-598
3. Gim, T. (1997), *Modelling and Evaluation of Time-varying Thermal Errors in Machine Tool Elements*, PhD Thesis, Cranfield University, Cranfield, Bedford, MK43 0AL, UK

Load Capacity of an Externally Pressurized Air Bearing Used in Measuring the Power Drawn in Stirred Tanks

David Palomino M.¹, Guillermo Robles M.¹, Gabriel Ascanio G.¹, Enrique Galindo F.²

¹Departamento de Diseño Mecánico, Centro de Instrumentos, UNAM.
A. P. 70-186, Ciudad Universitaria, México D. F. 04510, MEXICO.
Phone: +52 (5) 622 8654, Fax: +52 (5) 622-8620.

²Departamento de Bioingeniería, Instituto de Biotecnología, UNAM.
A. P. 510-3, Cuernavaca, Morelos, 62250. MEXICO.

Abstract

In order to predict the load capacity of an externally pressurized air bearing, an analysis of the air flow through the clearance formed between its components is carried out. The dimensionless Reynolds' equation which describes the movement of fluid in the bearing is solved numerically to obtain the pressure distribution neglecting the curvature effects and assuming isothermal compressible flow. The load capacity is also calculated by numerical integration over the total surface wetted by the air. An experimental device was constructed to carry out experimental measurements. The pressure supplied was measured at different loads and these results were compared with the numerical ones obtaining maximum differences of roughly 9%.

Introduction

The rotating machinery is used extensively in diverse applications as power stations, marine propulsion, aircraft engines, machine tools, automobiles, mixing systems, ultraprecision equipment, etc., and is supported generally by one or more bearings which play a vital role in determining the behavior of the rotating system under action of both static and dynamic loads. In order to design a system, it is necessary to know the types of bearings and the performance characteristics associated with each one of them. A type of bearing commonly used today is the air bearing that has many advantages as to reduce the power loss, to support relatively high load even without spin, are free of friction due to the air film generated by the compressed air, compared with oil bearing have less heat generation, low contamination and high accuracy [1,2]. This paper presents results of the load capacity in an externally pressurized air bearing obtained by means of the wetted surface and the pressure supplied.

Description

The bearing, sketched in fig. 1, consist of two components [3]. The static one is a dish provided with a conical cavity in its upper surface and a shaft passing through its center. The dynamic part is a conical shaped disc provided with a cylindrical hole in its center so that it can freely rotate around the shaft of the static component. Pressurized air is supplied in the static component, is expanded by nozzles discharging along of three shallow circular channels and, finally, is distributed on its surface. The channels allow the bearing to have a higher load capacity and help to distribute the air to all its parts. The air flows through the clearance existing between both

components of the bearing exerting a thrust force and keeping the dynamic part, together with its load, floating by means of an air cushion during the power drawn measurements. The air gets out finally to the atmosphere. The shaft of the static part have also nozzles that distribute air on its surface and both components have never any contact.

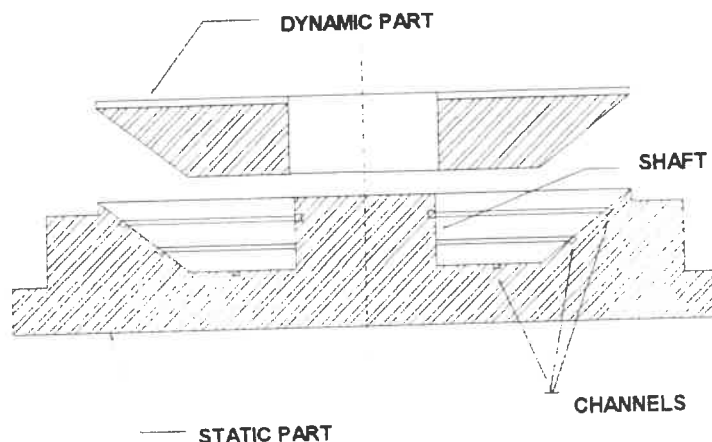


Fig. 1. Scheme of the pneumatic bearing

Theoretical Analysis

The governing equation which describes the characteristics of the pneumatic bearing can be obtained theoretically modeling the movement of fluid film. This equation was derived by Reynolds [4] for an isothermal compressible fluid and is presented as follows with the geometry shown on the fig. 2:

$$\frac{\partial}{\partial x} \left(\frac{\rho h^3}{\mu} \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial y} \left(\frac{\rho h^3}{\mu} \frac{\partial p}{\partial y} \right) = 6 \left[u_0 \frac{\partial}{\partial x} (\rho h) + v_0 \frac{\partial}{\partial y} (\rho h) \right] + 12 \frac{\partial}{\partial t} (\rho h) \quad (1)$$

where p , ρ , μ , and h are the pressure, density, viscosity and film thickness, respectively; and u_0 and v_0 are the velocities in the x and y directions, respectively.

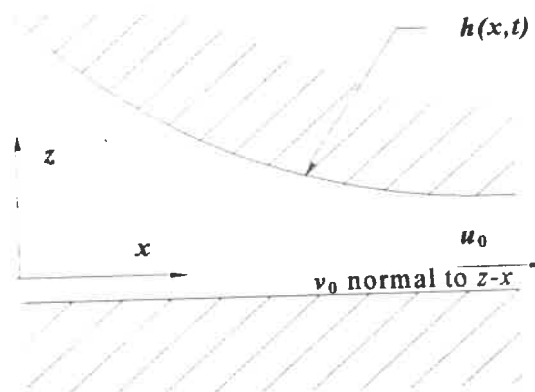


Fig. 2. Flow channel of lubricant

By introducing the ideal gas equation and assuming viscosity constant, steady state and static conditions, the Reynolds equation (1) is reduced to:

$$\frac{\partial}{\partial x} \left(ph^3 \frac{\partial P}{\partial x} \right) + \frac{\partial}{\partial y} \left(ph^3 \frac{\partial P}{\partial y} \right) = 0 \quad (2)$$

The discharging pressure p_o on the channels is obtained by means of the inlet pressure p_s and the adiabatic coefficient γ assuming a isentropic process [5]:

$$p_o = p_s \left(\frac{2}{\gamma + 1} \right)^{\gamma/(\gamma-1)} \quad (3)$$

The following dimensionless variables are used to generalize the solution, being p_a , L and c the atmospheric pressure, reference length and clearance, respectively [6].

$$P = \frac{p}{p_a}, \quad X = \frac{x}{L}, \quad Y = \frac{y}{L}, \quad H = \frac{h}{c} \quad (4)$$

By substituting in equation (2) it is transformed in:

$$\frac{\partial}{\partial X} \left(PH^3 \frac{\partial P}{\partial X} \right) + \frac{\partial}{\partial Y} \left(PH^3 \frac{\partial P}{\partial Y} \right) = 0 \quad (5)$$

The boundary conditions are $P=P_o/p_a$ at each outlet of nozzles or channel, $P=1$ at the outlet to the atmosphere, and $P(X)=P(X+2\pi)$ as a periodicity condition.

The load capacity supported by the bearing is obtained by integrating the pressure distribution over the total surface A wetted by the air:

$$\bar{F} = p_a \int P dA \quad (6)$$

A computational code was elaborated by using an iterative scheme of finite differences where the value of pressure at each node is determined successively up to satisfy a given tolerance. These results are shown in fig. 3 where it can be observed an exponential growth.

Experimental Measurements

Experimental measurements to determinate the static load supported by the bearing were carried out. Pressurized air at distinct pressures was supplied up to lift up the load together with the dynamic part. A tangential force was applied to the dynamic part to permit free spin and to verify the floating. Pressure measurements were taken with a Bourdon manometer up to raise each load. These results are shown in fig. 3, together with the numerical ones presenting the same behaviour.

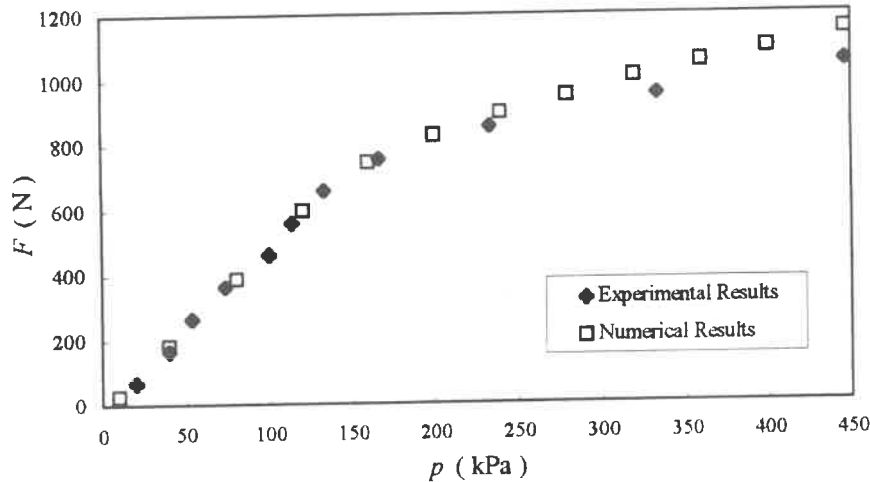


Fig. 3. Comparison between calculate and experimental results

Conclusions

The numerical calculations and experimental measurements present differences lesser than 9%. These differences are due to the limitations in the static model and to the nonuniformity of the load applied that causes instabilities and friction between the static and dynamic parts. A dynamic model will be solved in the future to obtain the dynamic characteristics and a better approximation between experimental and calculated results.

Thus, numerical and experimental load capacity results in an externally pressurized pneumatic bearing were presented in this paper with an acceptable approximation making a good prediction of the bearing performance.

Acknowledgment

This project was partially supported by funds from DGAPA-UNAM grant No. IN-503295.

References

1. Gross, W. A., *Fluid Film Lubrication*, New York, Wiley (1980).
2. Goodwin, M. J., *Dynamics of Rotor-Bearing System*, Unwin Hyman LTD, London (1989).
3. Reséndiz, R., Martínez, A., Ascanio, G., Galindo, E., "A new pneumatic bearing dynamometer for power input measurement in stirred tanks", *Chem. Eng. Tech.*, **14**, (1991). 105-108.
4. Reynolds, O., "On the theory of lubrication and its application to Mr. Beauchamp Tower's experiments, including an experimental determination of the viscosity of olive oil". *Phil. Trans. Roy. Soc. London*, **177**, pt. 1, (1886). 157-234.
5. Currie, I. G., *Fundamental Mechanics of Fluids*, Mc Graw-Hill Book Company, USA, 1974.
6. Han, D., Park, S., Kim, W., Kim, J., "A study on the characteristics of externally pressurized air bearings", *Precision Eng.*, **16**, (1994). 164-173.

Design of an Ultra Low Speed Straight Motion System Based on a High Viscosity Fluid Bearing

J.W. Spronck, P.F.A.M. van Putten and G.J.P. Nijse

Delft University of Technology, Laboratory for Micro Engineering,
Mekelweg 2, 2628 CD Delft, the Netherlands

Tel. +31-15- 2781824; fax +31-15-2781887, E-mail J.W.Spronck@WbMT.TUdelft.NL

Abstract

An ultra low speed straight motion system is described. The system consists of a fine stage on top of a coarse stage. The coarse stage consists of a moving beam, a base plate and a side plate, separated by a thin grease film. The coarse stage must realize a very smooth motion within the control capabilities of the fine stage. The measurement results show that the proposed system can be used to establish a very constant low speed motion. More research has to be done to optimize the coarse stage.

Introduction

How to design a simple nanometer straight motion system that provides for a constant velocity in the range from $1 \mu\text{m/s}$ to 1 mm/s , over a travel range of approximately 100 mm?

The mechanical setup can be described as a fine stage on top of a coarse stage.

The coarse stage must realize a very smooth linear motion within the control-capabilities of the fine stage.

The fine stage (not drawn or discussed in this paper) has been equipped with piezo-actuators to realize a fast tracking relative to a straightness reference standard. The range of the fine stage is a few micrometers.

This paper describes the mechanical design and the first measurement results of the coarse stage. The working principle of the coarse stage is shown in figure 1.

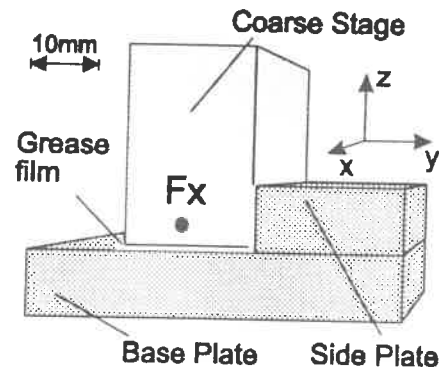


Fig. 1. The coarse stage is supported by a grease bearing on a base plate. F_x is the driving force in the x-direction.

1. The design concept of the coarse stage

The design of the coarse stage has been simplified to a great extent by the multifunctional use of grease. We used grease named *Kilopoise 0001* (a product of Rocol Ltd., UK). The film thickness between the stage and the base plate is approximately 20 micrometer.

The most important advantages of grease in relation to the motion of the stage are:

1. The damping characteristic of the grease lubrication provides for a simple linear relation between the driving force F_x and velocity V_x ;
2. High frequency deviations from the ideal momentary position and high frequency straightness errors that are normally caused by surface roughness and vibrations of the bearing, are eliminated to a great extent by the grease.

The most important advantages of grease in relation to the bearing of the stage are:

1. A very low cost lubricant and bearing with a high damping ratio is achieved;
2. The bearing provides high stiffness at high frequencies and almost zero stiffness at very low frequencies;
3. Easy and low cost assembly of the bearing.
4. No special requirements are needed for the bearing surfaces.

The overall advantages of the bearing are:

1. Very compact design;
2. Low cost;
3. Flexible design

The overall disadvantages of the bearing are:

1. The thickness of the grease-layer is not critical but must be controlled within certain limits. Stick slip and the roughness of the bearing surfaces affect the smooth motion.
2. Grease is a non-Newtonian fluid. The apparent viscosity depends on the shear rate and shear amplitude.
3. The viscosity of the grease depends on the temperature.

2. Physical model

The rheological behavior of greases in a lubricated bearing is a complex problem. Grease is a non-Newtonian colloidal system. When subjected to oscillatory shearing of a certain amplitude, the grease exhibits a non-linear response in the output signal. The *Kilopoise 0001* grease used exhibits a high apparent viscosity at low shear rates and a low apparent viscosity at high shear rates. The viscosity also depends on the shear amplitude. The exact behavior has to be determined experimentally.

A linear model of the bearing has been derived to predict the behavior of the bearing.

The bearing can be described as a damper. The differential equation of a mass-damper system is:

$$m \cdot \ddot{x} + c \cdot \dot{x} = Fx,$$

where x is the displacement of the stage, Fx is the force applied to the stage, m is the mass of the stage and c is the damping constant. The damping constant depends on the geometry of the fluid film and the viscosity that is assumed to be constant in this idealized model.

$$c = \frac{l \cdot w}{h} \cdot \eta.$$

Here, l is the bearing length, w is the bearing width; h is the fluid film thickness and η is the coefficient of viscosity.

The viscosity of the grease used is 0,25 Kilopoise (in SI units: 25 kg/ms) at 25° C. The linear model of the bearing behaves like a low-pass filter.

3. The experimental setup and results

Experiments have been performed to obtain information on the following:

1. What is the relation between the driving force and the velocity of the stage?
2. How smooth is the displacement in x -direction? This should be within the control capabilities of the piezo-driven fine stage.
3. How smooth are the straightness errors? For the same reason as mentioned above they must be within a few micrometers.

A laser interferometer was used for all displacement, velocity and straightness measurements.

3.1 Translation measurement; force-velocity

The experimental setup for the translation measurement is shown in figure 2.

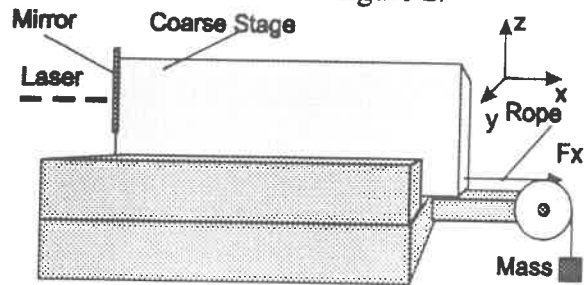


Fig. 2. Scheme of the experimental setup for measurements in the x-direction. A steel slider placed against two steel beams, separated by a thin film (20um) of grease. The force is applied by a variable mass hanging on a rope

During the very first experiments, the stage was pushed by hand. The results are represented in figure 3 and 4.

Table 1. The average Speed(AVG dx/dt in $\mu m/s$) by applying different forces. The relation between speed and force is linear within 10%.

Force	0	1,8	2,8	3,8	4,8	5,8
AVG dx/dt	0	76,5	128	180	228	282

In figure 5 we see the response of the stage when the driving-force is changed instantly from 2,8 N to 0 N. In the experimental setup a force in the range from 1,8 to 5,8 N gave a velocity from 76,5 to 282 $\mu m/s$. Results are shown in table 1.

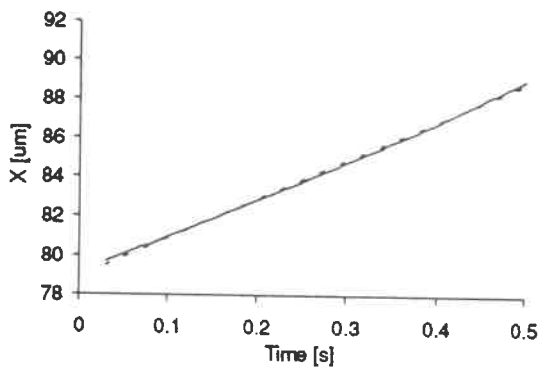


Fig. 3. Hand-driven x translation (from 80 to 140 μm in 2.5 seconds) and ideal position line (dashed). The deviation from the ideal position is less than 1 μm .

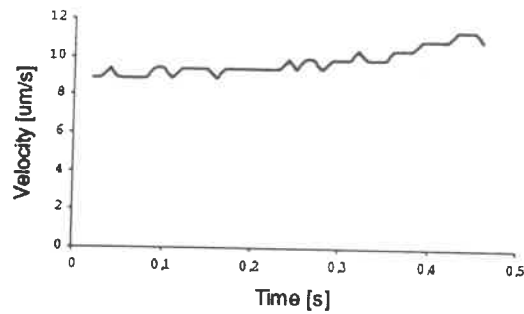


Fig. 4. Hand driven velocity vs. time for the first 0.5 seconds. The velocity is about 10 $\mu m/s$. The velocity error is less than 2 $\mu m/s$. The high frequent changes are due to the limited resolution of the laser.

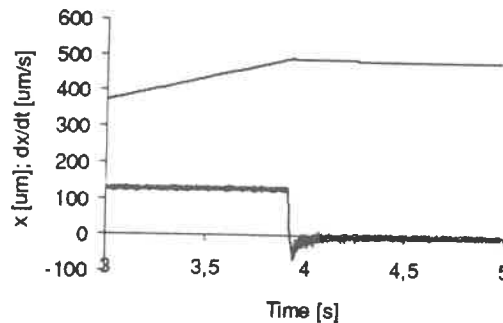


Fig. 5. A step in the applied force from 2,8 to 0 Newton. The step is generated by lifting the weight instantly.

3.2 Translation measurement: straightness

The experimental set-up for the straightness measurements is shown below.

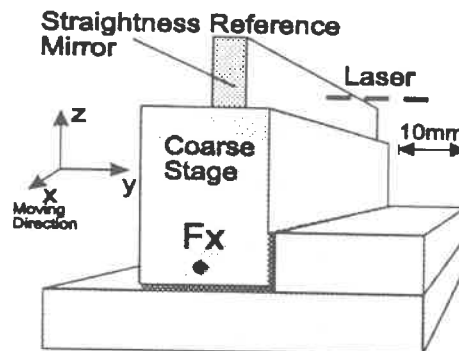


Fig. 6. Straightness measurements setup. A zerodur straightness reference mirror is placed on top of the stage. The distance between stage and laserinterferometer is measured while moving the stage.

During 120 seconds, the movements in the y-direction have been measured. The stage has moved over a distance of 9 mm in the x-direction.

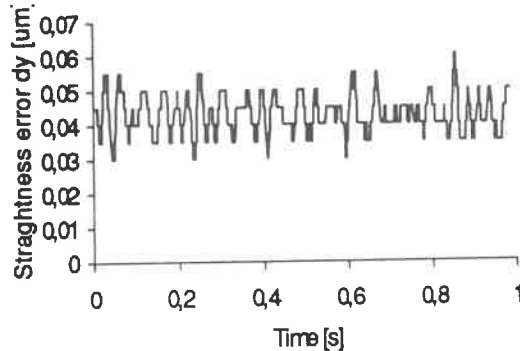


Fig. 7. Plot of the straightness error during the first second. The speed in the x-direction is 75 $\mu\text{m/s}$. The straightness error is about 30 nm over a distance of 75 μm .

The straightness error over the distance of 9 mm is 0,14 μm .

The behavior of the stage in the z-direction is assumed to be similar.

3.3 Repeatability and drift

The behavior of the bearing is not constant over a few days. The film thickness decreases slowly to a constant value. Different loads on top of the bearing also influence the film thickness. The film thickness should be controlled to solve this problem. Another problem is that the temperature of the grease raises due to internal friction. This results in a lower constant of viscosity, which influences the relation between the velocity and the applied force.

When the driving force is turned off, the velocity drops to zero instantly and seems to stay zero. Position measurement showed that the stage is moving 2 μm in 10 minutes.

4. Comparison of the measurement results with the physical model.

From the measurement results we can conclude that the bearing behaves as a low-pass filter.

The stiffness of the bearing in the x-, y- and z-direction is high for high frequencies and low for low frequencies.

For small steps in the applied load, the physical model is satisfying. High steps could show the non-linear behavior, but are not needed in this application.

For a specific velocity range, one should measure the bearing characteristics experimentally to derive the damping constant.

5. Conclusions

The performance of the course stage is very promising. One should be aware that no special construction tools or materials were used.

1. The relation between the driving force and the velocity of the stage is linear within 10 %.
2. The smoothness of the displacement in the x-direction and the straightness errors are within tenths of nanometers.
3. More research has to be done to derive a non-linear physical model for grease-film bearings.

The piëzo driven top stage is under development and is expected to compensate the remaining tracking errors.

References

- [1] A. Cameron, *The Principles of Lubrication*. Longmans, London, 1966
- [2] B. Pugh, *Practical Lubrication*. Newnes-Butterworth, London, 1970

COMPARISON OF CARBON FIBER ALUMINA CERAMIC AND ALUMINUM BEAMS

Drew Devitt, Nick Sotiropoulos, Josh Radice
New Way Machine Components, Inc.
Aston, PA 19014

Carbon fiber composites have made a dramatic influence in the space and aero space industries. Lower weight, high strength, stiffness and controllable coefficient of thermal expansion of carbon fiber are several of the technical properties that could be employed with advantage in precision structures for metrology or manufacturing applications. The purpose of this paper is to help evaluate this potential.

To this end, we acquired an aluminum beam, a ceramic beam, and manufactured a carbon fiber beam in our shop using resin transfer molding. It should be noted that the aluminum beam had no machining on its surfaces and was not anodized. The composite beam was designed with unidirectional fiber, balanced with the resin to have near zero coefficient of thermal expansion. All three beams were 2.5 inches square with a .250 inch wall thickness and 21.5 inch length.

Three basic tests were performed on the beams: bending strength, thermal distortion, and dynamic analysis. The tests and their results are described below.

	BEAM MASS (gms)	DYNAMIC CHARACTERISTICS		THERMAL DISTORTIONS		BENDING STRENGTH	
		FIRST NATURAL FREQUENCY	DAMPING FUNCTION	AXIAL GROWTH (micro-inch)	BENDING (sensor C) (micro-inch)	DEFLECTION (2 1/2 lb. LOAD) (micro-inch)	DEFLECTION (5 lb. LOAD) (micro-inch)
CERAMIC	3495	3350	.12%	47	182	131	261
ALUMINUM	2110	1295	.20%	201	133	413	829
CARBON FIBER	1135	1275	1.4%	88	687	637	1273

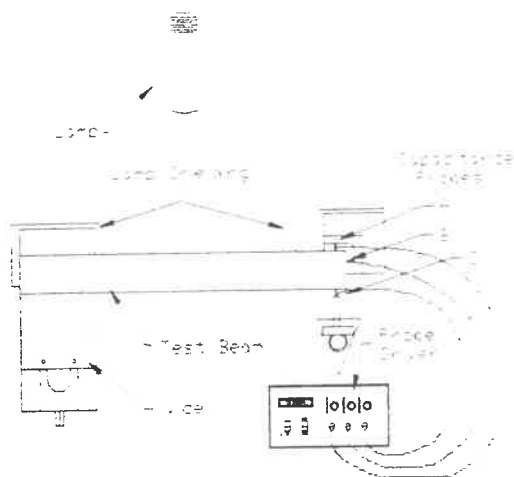
BENDING STRENGTH

The test beams were held in a vise 5 inches wide resulting in a cantilever of 16.5 inches. Two and a half pound and five pound loads were applied to the end of the beam. Measurements were taken using a capacitance probe and a steel gage block attached to the underside of the beam. Results are listed on the right side of the above summary table. The carbon fiber deflected the most and also took a noticeably longer time to return to zero after the loads had been removed. This lag may be attributed to micro deformations of the epoxy along the fiber interface which may also be responsible for the excellent damping. Unfortunately it is not advantageous in this application. It should be noted that this composite was designed for low coefficient of thermal expansion, not for high stiffness.

THERMAL DISTORTION TEST

Deflection in micro-inches

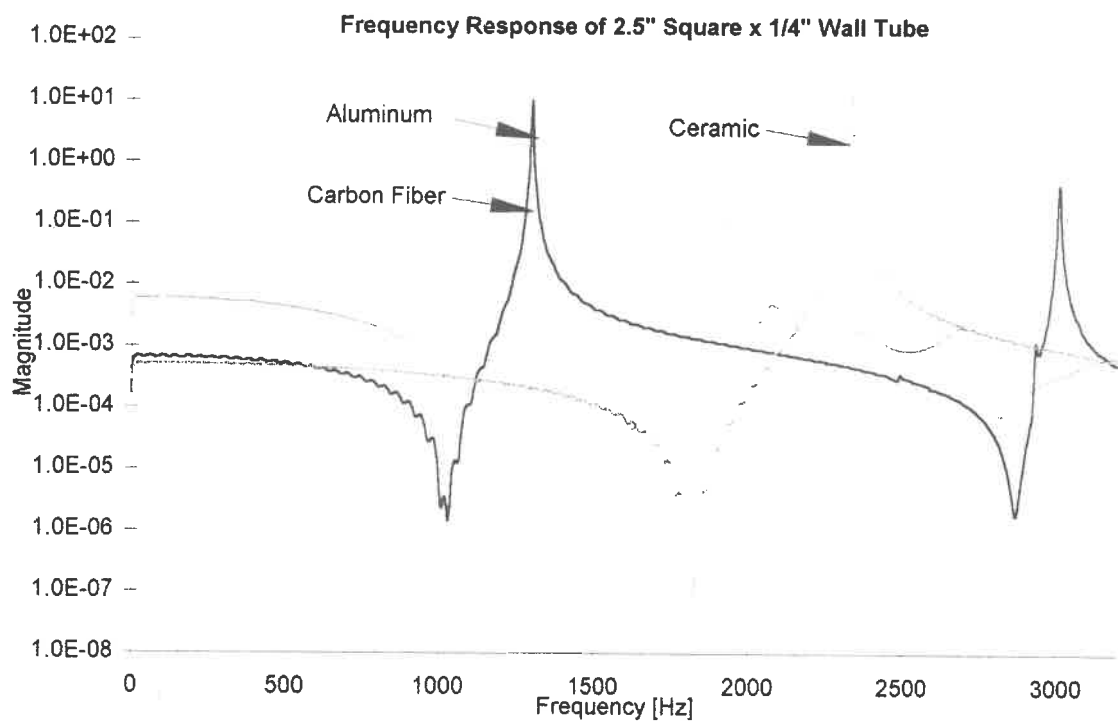
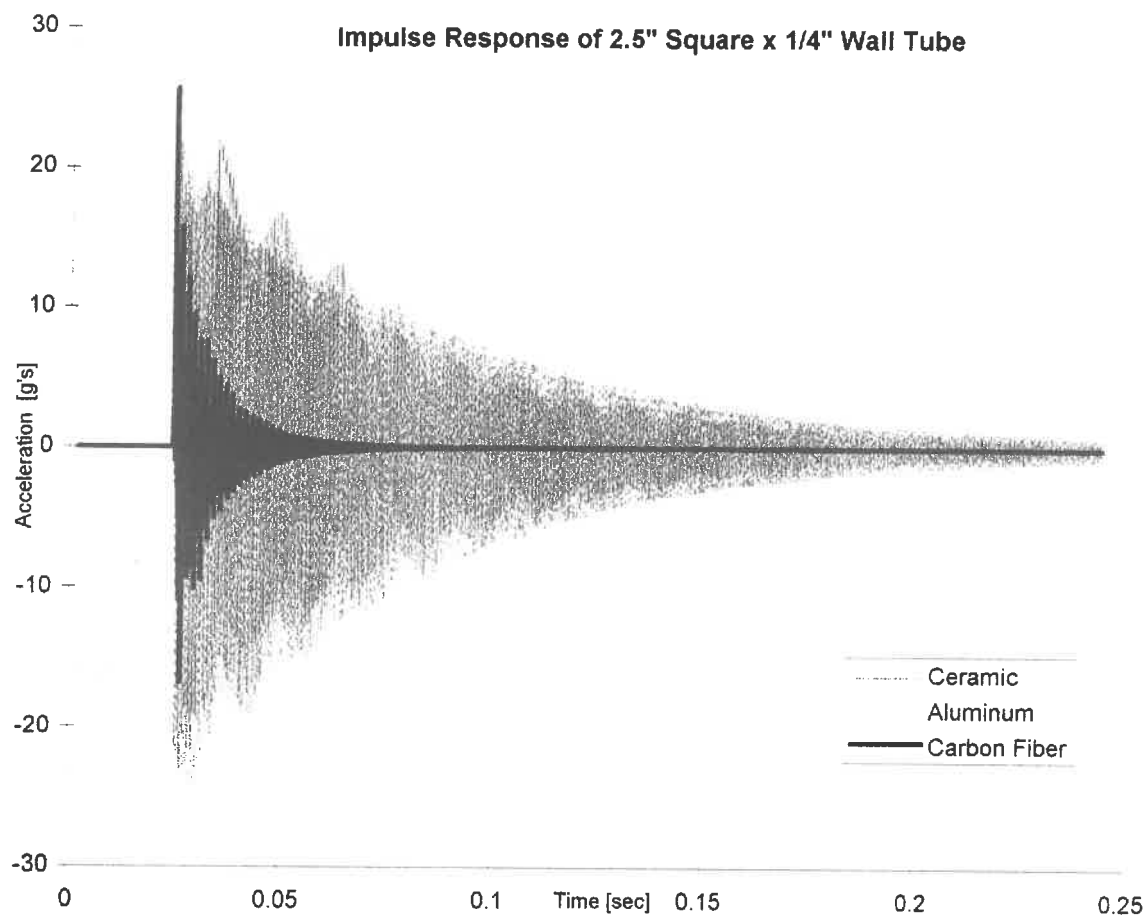
BEAM	SENSOR	LIGHT ON 3 minutes	LIGHT OFF 3 minutes	LIGHT OFF 6 minutes	LIGHT OFF 9 minutes
CERAMIC	A	70	70	35	13
	B	47	44	37	34
	C	182	82	30	16
ALUMINUM	A	-97	48	43	39
	B	201	212	207	205
	C	133	1	11	17
CARBON FIBER	A	-608	-314	-218	-147
	B	88	61	48	39
	C	687	394	291	231



This test was performed on a large granite surface plate. Five inches of each tube was held in a vice and steel gage blocks were attached to the rams as indicated in the diagram to provide a conductive surface for the capacitance probes. The capacitance probes measured displacements caused by thermal distortions. The probes and vise were shielded from a 200 watt light placed 12 inches from the surface of the beam. In all three cases dimensional distortions were evident within 5 seconds of exposure to the light. It is likely that the black carbon fiber

absorbed more energy than the white ceramic or reflective aluminum. It should be noted that in practice aluminum is hard coated or anodized dark gray or black. Thus, a more representative test would have been to use an aluminum beam treated in this manner because of the specific interest in determining actual responses to radiant heat inputs. It also would have been interesting to have outfitted the beams with temperature sensors as the heat conductivity of the materials certainly affected in the distortion patterns.

The carbon fiber beam showed a low axial growth but exhibited a warp much larger than expected. The ceramic beam also warped more than expected but was the most stable in the axial direction. The ceramic beam was the easiest to maintain good reference zero's when testing. While testing the aluminum beam care had to be taken to keep body heat or breath away from the beam as they would have an immediate dimensional effect. The conductivity of aluminum appeared to help minimize warping by equalizing temperature gradients quickly.



DYNAMIC ANALYSIS

The dynamic characteristics of the three beams were measured using an impact hammer, an accelerometer, and a dynamic signal analyzer. Each beam was hung from surgical tubing and excited using the impact hammer. The resulting accelerometer response was measured. From these tests, the natural frequencies and damping can be evaluated.

The first natural frequency of the carbon fiber was almost the same as the aluminum, although its second natural frequency was much lower. We suspect this was due to an out of square or diaphragm type vibration of the side wall. The unidirectional fiber orientation was selected for simplicity of obtaining a low coefficient of thermal expansion in the axial direction. It is expected that bi-directional or multi-directional fiber orientations will improve the second and higher natural frequencies. A higher fiber content combined with bi-directional or multi-directional components should significantly increase static stiffness and the first natural frequency.

The first natural frequency of the ceramic was significantly higher than both the aluminum and the carbon fiber. As expected though, the ceramic had the lowest damping function. The carbon fiber damping function was very high with vibration decays approximately seven times faster than the aluminum. This should result in faster setting times and excellent dynamic performance.

SUMMARY

Carbon fiber shows promise in large part to its light weight and high damping. If increased thermal stability and bending stiffness can be achieved, carbon fiber could become an attractive design material for precision engineers.

We recognize moisture absorption as another potential source of distortion for carbon fiber and so we will have information on stabilizing carbon fiber against moisture absorption on our poster.

Our thanks for their contributions to our project:

Dr. Eric Marsh
Dept. of Mech.Engrg.
Penn State University

and

Dr. Suresh G. Advani
Dept. of Mech.Engrg.
Center for Composite Mfg.
Univ. of Delaware

Estimation of Thermal Contact Resistance by Electrical Contact Resistance Measurement

Kiyoshi Mizuhara and Norimitsu Ozawa
Mechanical Engineering Laboratory, AIST, MITI
1-2 Namiki, Tsukuba, Ibaraki, 305 JAPAN

1. Introduction

Thermal deformation of machine tools is very important problem for the working accuracy. To solve this problem, it is necessary to analysis of temperature distribution and thermal displacement of the machine tool structure. Particularly, study of the thermal behavior on contact surfaces of the machine tool is essential. In general, there are two kinds of contact surfaces between structural components of machine tool; two components are fixed by bolt - nut or any clamping device, and two components are sliding with each other. In this paper, we discuss about the former contact surface and a method to evaluate the characteristics of the thermal contact resistance (hereinafter called the thermal resistance) on the contact surfaces has been studied. Since the thermal resistance is greatly influenced the contact conditions of contact surfaces, the electrical contact resistance (hereinafter called the electrical resistance) has been used to investigate the contact conditions. Several factors are affecting the electrical resistance on contact surfaces e.g. (1) surface roughness, (2) waviness of contact surfaces, (3) contact pressure between contact surfaces, and (4) substances existing on contact surfaces. As these factors also influence on the thermal resistance on contact surfaces, it is assumed that a correlation exists between the thermal resistance and the electrical resistance on contact surfaces. A number of research works not only the thermal resistance but also the electrical resistance have been carried out,^{(1), (2), (3), (4)} however, there are no research on the correlation between two resistance. From the point of view, this paper first describes the results of experiment of the electrical resistance on contact surfaces and, then, the results of the thermal resistance measurement. Finally, correlation between the electrical resistance and the thermal resistance will be discussed.

2. Experimental set-up and procedure

Figure 1 shows the experimental set-up for the thermal resistance measurement and test specimens. For the observation of changes in the electrical and thermal resistance caused by contact pressure, a standard pressure indicator with nominal weight and the hydraulic power loading device were used in increasing/decreasing the contact pressure, and the applied load was detected by a load-cell and indicator. As shown in this figure, the test specimens are cylindrical form (50 mm in diameter and 50 mm long) and end surfaces of each specimen are machined to serve as contact surfaces. Three types of contact surface were used: lapped surface (steel), ground surface (steel), and scraped surface (cast iron). Profile of the contact surface of test specimens is shown in Figure 2. As presented the figures, ground surface is the best surface waviness and lapped surface has the best surface roughness. To investigate the effects of lubricant, measurement of resistance were carried out under the following two condition: DRY (the contact surfaces were washed by benzine) and WET (oil was injected between contact surfaces washed by benzine). The total initial loading of 64N on the contact surfaces included weights of the test specimen, the heat source, the heating block and the load cell. Consequently, initial contact pressure was 0.03MPa. All experiments were carried out in a temperature controlled room within 20 ± 0.5 degrees.

In case of the thermal resistance measurement, as shown in Figure 1, heat source is heated using a heating block provided with four thermoelements. The heat flow from the heat source passes through the contact surface between two specimens. Temperature of the heat source was controlled by means of regulating the DC power supply in accordance with the output of the thermal recorder. The thermal recorder is monitored the heat source temperature measured by thermistors embedded in the heat source. The thermal resistance on the contact surfaces was measured using copper-constantan thermocouples inserted into four holes (S_1 , S_2 , S_3 , and S_4) of the specimens. And output signals of thermocouples were transmitted via digital thermometers to an analog computer for calculation using the extrapolation method. Since the electrical resistance to be measured is very small value, thus, we used a double bridge. Current and voltage terminals for the double bridge were connected to the same holes (S_1 , S_2 , S_3 , and S_4) of the specimens instead of the thermocouples. Since the electrical resistance is not stable immediately after loading, measurements were made after the output signal saturated. The electrical resistance was measured while the load applied on the contact surfaces were

increased from 0N to 2942N (contact pressure 1.53MPa) in increments 490N, and also unloaded by the same increments. Three consecutive measurements as described above were conducted for each contact surface.

3. Experimental results and discussions

3.1 The electrical resistance Figure 3 and 4 represented the experimental results of relationship between the electrical resistance and contact pressure under loading in DRY and WET conditions of ground specimen. The electrical resistance decreased with increasing the contact pressure. The results clearly demonstrated that expansion in real contact area due to the contact pressure caused reduction of the electrical resistance. Value of the electrical resistance in WET contact condition, Figure 4, is about 1/5 of DRY contact condition shown in Figure 3. Various researches have been carried out concerning the electrical resistance of thin oil film.^{4), 5), 6)} These researches indicate that the electrical resistance of oil film thickness equal to the oil molecule became to approximately zero based on the tunnel effect. It is evident that the electrical resistance reduction on WET condition is caused by the tunnel effect. In addition, the electrical resistance increased with increasing frequency of load application, and contact pressure at 1.53 MPa it becomes about 5 times the value taken from the first measurement. This phenomenon are interpreted by the fact that: (1) current flowing through a very small contact area during measurement, (2) microscopic electric discharge, (3) contact parts suffer change in quality (such as oxidation), and (4) as result of formation of certain material between the substrate metal and the contact parts. Figure 5 show the effect of finished surface on the electrical resistance under DRY contact condition. As this figure shows, the rougher the surface, the larger the electrical resistance, and more susceptible to the contact pressure.

3.2 The thermal resistance Figure 6 presents relationship between the thermal resistance and the contact pressure under DRY and WET contact conditions. Effect of the contact pressure and the lubricant are similar to the experimental results of the electrical resistance, however variation caused by the frequency of load application is obviously small values. Expansion of real contact area and existence of the oil cause the reduction of the thermal resistance. On the other hand, for the analysis of the different behavior between the electrical and thermal resistance on the load application frequency, further investigation is necessary. Figure 7 shows the relationship between the thermal resistance and the finishing methods. Behavior of the thermal resistance is approximately same in the case of the electrical resistance shown in Figure 5.

4. Correlation between the thermal resistance and the electrical resistance

Figure 8 shows the relationship between the thermal and electrical resistance for DRY and WET conditions at contact pressures of 0.53MPa and 1.53 MPa respectively. As is clear from the figure, the thermal and electrical resistance can be represented by a linear relationship on log-log curves independent of contact pressure, surface roughness and contact conditions (DRY and WET). In other words, it has been verified by experiments that the relationship between the thermal and electrical resistance changes exponentially. By representing the thermal and electrical resistance in terms of R_T and R_E respectively, the thermal resistance R_T is given by the following equation,

$$R_T = 0.35 R_E^{0.74}$$

5. Conclusions

From the measurements of the thermal and electrical resistance under the same conditions, we led to the following conclusions:

- (1) The electrical resistance is small value for combinations of test specimens having less surface roughness, and tends to decrease with increasing contact pressure. The rate of change in the resistance value against contact pressure decreases for combinations of test specimens having smooth surfaces. For WET conditions, the electrical resistance is smaller, and is less influenced by contact pressure and surface roughness than in DRY conditions.
- (2) In case of the thermal resistance, effect of the contact pressure and the lubricant exhibited a tendency similar to the electrical resistance.
- (3) The resistance due to the frequency of load application represented obvious difference between the electrical and thermal resistance. Analysis of the phenomenon is a future subject.
- (4) The relationship between the thermal and electrical resistance is exponential, irrespective of contact and surface conditions. Therefore, measurement of the electrical resistance, which is easy to carry out, makes it possible to estimate the thermal resistance.

References

- 1) Weills, N. D. and Ryder, E. A., Trans. ASME 71(1949)259
- 2) Kouwenhove, W. B. and Potter, J. H., Weld. J., 27(1948)515
- 3) Uppal, A. H. and Probert, S. D., Wear., 15(1970)271
- 4) Holm, R., Electrical Contact, Springer-Verlag(1967)
- 5) Tamai, Y., J. of JSLE, 11-3(1957)115, (in Japanese)
- 6) Wilson, R. H., Proc. Roy. Soc., A212(1952)450

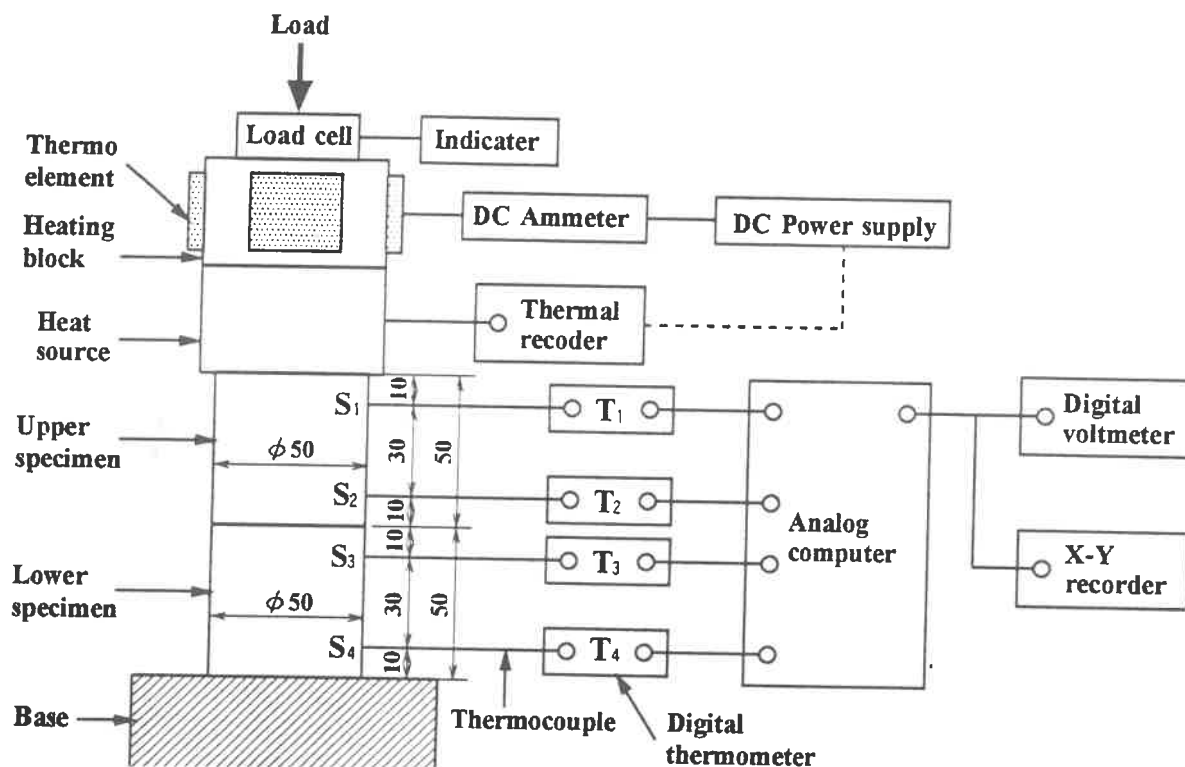


Fig.1 Experimental set-up and test specimens

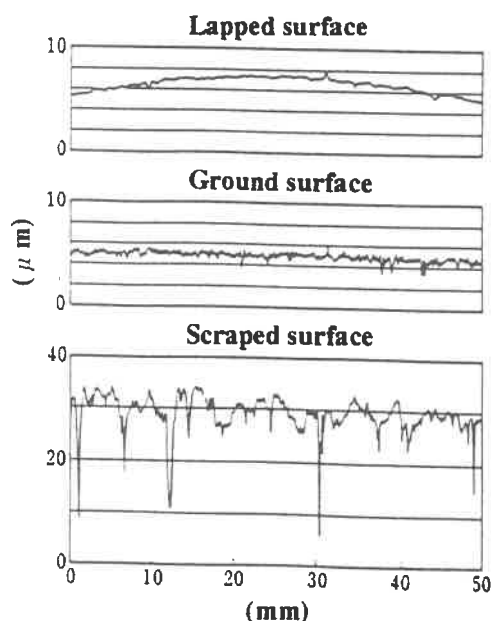


Fig.2 Profile of test specimens

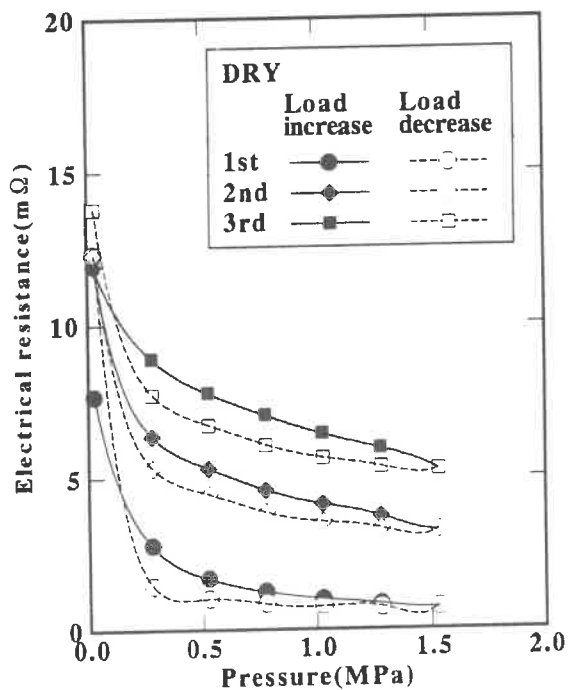


Fig.3 Relationship between pressure and electrical resistance (DRY) (Ground - Ground surface)

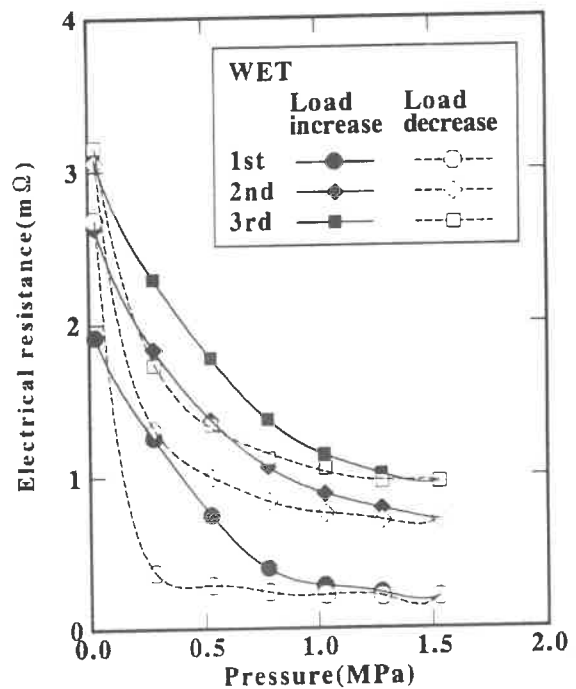


Fig.4 Relationship between pressure and electrical resistance (WET) (Ground - Ground surface)

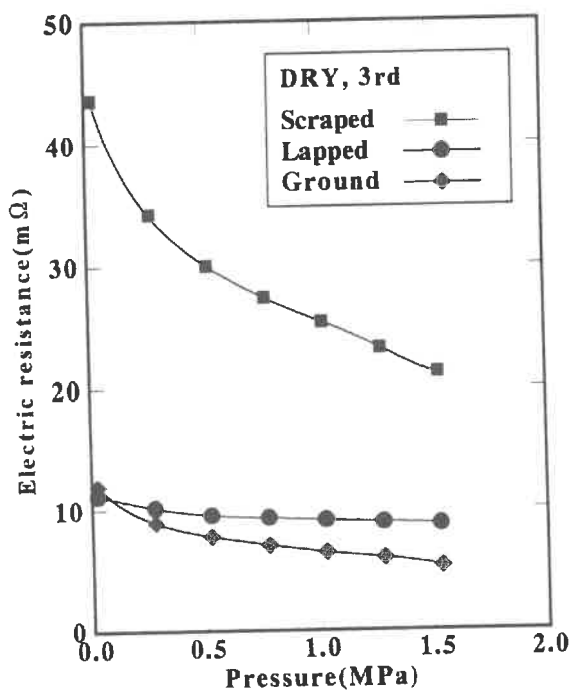


Fig.5 Relationship between finished surface and electrical resistance (DRY, 3rd)

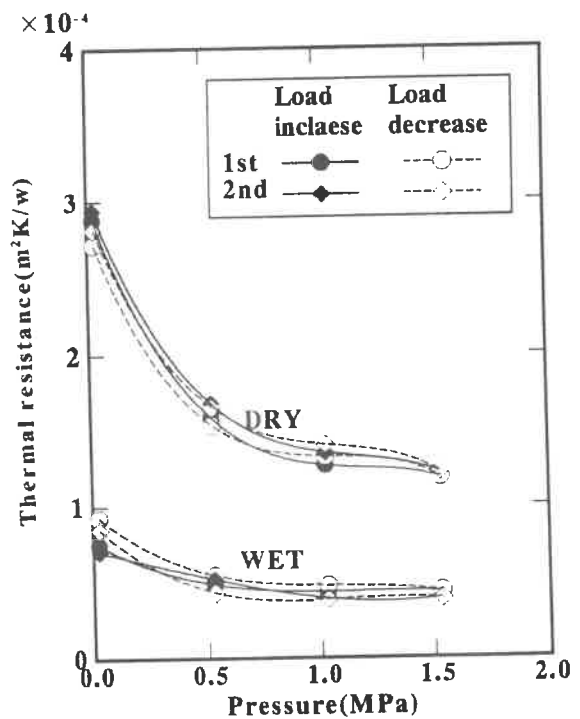


Fig.6 Relationship between pressure and thermal resistance (DRY and WET) (Ground - Ground surface)

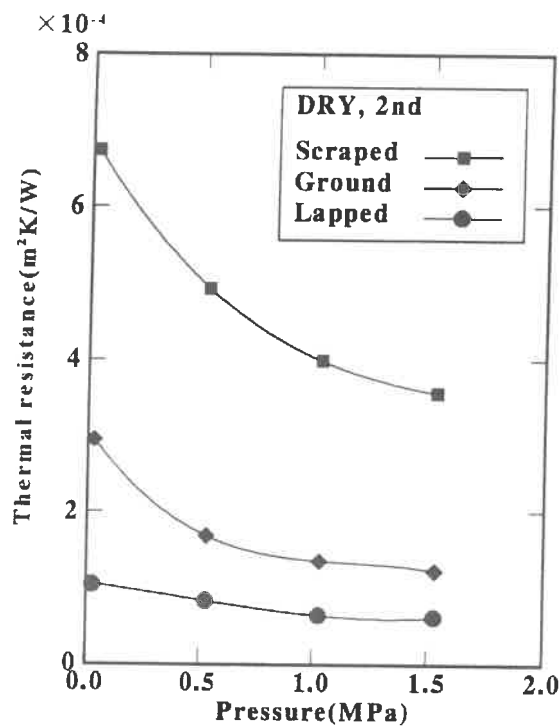


Fig.7 Relationship between finished surface and thermal resistance (DRY)
(Ground - Ground surface)

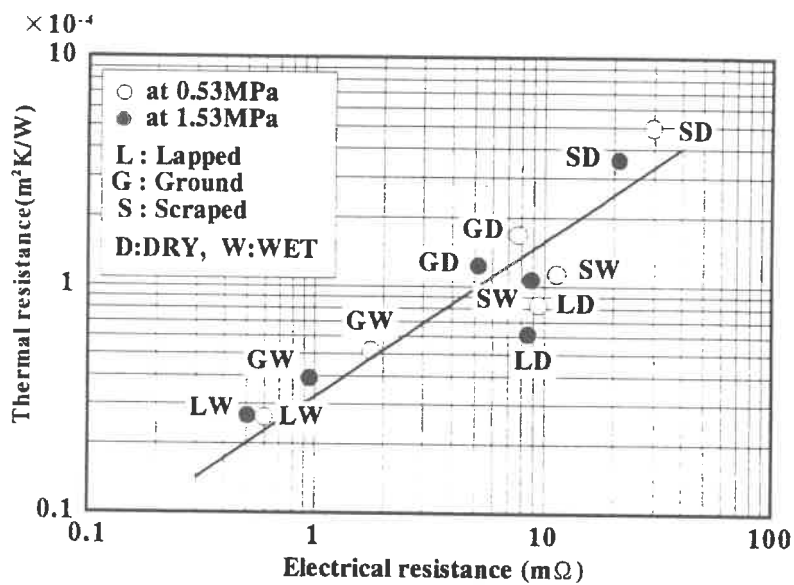


Fig.8 Relationship between electrical and thermal resistance

DESIGN OF SECOND GENERATION NANOMETRIC CUTTING INSTRUMENT

Robert Hocken, Ramesh Namperumal, Paulo Pereira, John Lovingood, Robert Edgeworth, Aditya Sharma, Michael Holmes, Balasubramaniam Muralikrishnan.
Center for Precision Metrology
University of North Carolina at Charlotte, NC -28223

INTRODUCTION

There is an ever increasing demand for manufacturing to dimensions in the nanometric range with subatomic level smoothness without residual stress and metallurgical defects. To achieve this, Single Point Diamond Turning (SPDT) can be used [1][2].

SPDT machines which can provide sub-micrometer level accuracy are commercially available. These instruments have been used successively to study the ultra-precision machinability of copper [2], mechanism of micro-chip formation [3], mechanical energy involved in micro-machining [4]. But the study of machining in the nanometric range requires a new generation of instruments since it is beyond the control range achievable by commercially available diamond turning machines. A first generation instrument which can achieve cutting at nanometric range was built at the Center for Precision Metrology at UNCC [5]. A second generation instrument was designed and built as a team project for a graduate-level instrument design course. This report describes the design of the second generation nanometric cutting instrument developed to investigate the mechanics of nanometric cutting.

CONCEPTUAL DESIGN

The basic principle of operation of the instrument is a single point diamond planer as shown in the figure. The cutting movement is provided to the sample mounted at the end of piezoelectric (PZT) tube, with the tool being held stationary. The infeed and crossfeed cutting motions are provided by the axial and lateral displacement of the piezoelectric tube scanner respectively. The feed motions are measured using a capacitance probe mounted to the PZT tube scanner. In order to study the mechanics of micromachining the cutting and thrust force are measured using two single axis force transducers attached to the tool holder.

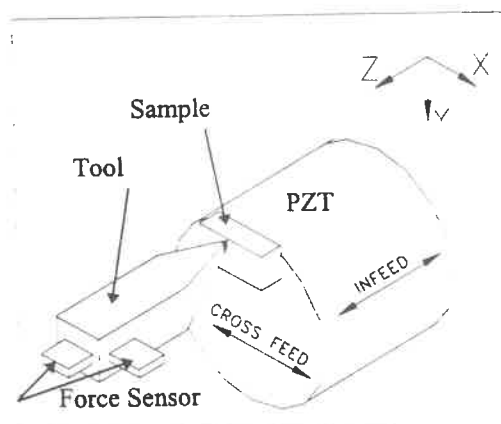


Figure 1 Principle of Operation

General considerations for the development of the instrument are dictated by various parameters depending on the intended functions. For example the instrument (1) should provide ability to measure forces in milli-newton range [1], (2) should provide a maximum of $2\mu\text{m}$ depth of cut and a maximum of $10\mu\text{m}$ length of cut, (3) should have

measurement resolution of cutting motions in the order of 0.1nm using a capacitance probe, (4) should have a stiffness of 1.8×10^7 N/m (10^5 lb./in), a typical value for the cutting instruments of this type [1], (5) should be completely vacuum compatible with all the components composed entirely of non-magnetic materials to allow observation within Scanning Electron Microscopy (SEM), (6) should be thermally stable for usage at normal room temperature environments, (7) overall outer dimensions should be less than 8 inches in diameter with center at the observation point (tool tip) accessible for observations inside a SEM.

INSTRUMENT DESIGN

To meet the above design requirement the nanometric cutting instrument, NanocutII as shown in figure 2 has been designed and built. The instrument has four modules namely:

- (1) The sample holder and sample mount,
- (2) PZT and capacitance probe assembly,
- (3) Tool holder and force sensor assembly,
- (4) Fine positioning system.

SAMPLE HOLDER AND SAMPLE MOUNT

This module consists of a sample holder and a sample mount or kinematic mount. The workpiece which is 5x5x10mm is glued at the top of the sample holder as shown. The sample holder made of Aluminum is provided with angular flexure by cutting a slit. This flexure is designed to provide a tilting adjustment up to 10^{-5} radians for the sample along X axis. Tilting arrangement enable the face of sample to be aligned with PZT axis during cutting tests. The tilting can be provided by adjusting the two alignment screws provided. The function of the sample mount is to provide a repeatable mounting of the sample onto the PZT assembly. The sample mount is designed using two flats (of hot pressed SiC), a V groove (of hot pressed SiC) and three SiC balls of 6.35mm (1/4") diameter. The preload, calculated to meet the minimum stiffness requirement, was 5N on each of the balls. Such preload force was achieved by three O rings arranged in such a way to provide easy access while assembling.

PZT AND CAPACITANCE PROBE ASSEMBLY

The PZT tube scanner mounted at the back of sample mount is used to provide cutting motions for the sample. By controlling the voltage applied to the quadrants of the PZT tube, displacements in two orthogonal directions can be achieved. A PZT tube suitable for providing 2μm depth of cut (Z) and 10μm length of cut (X) was chosen for

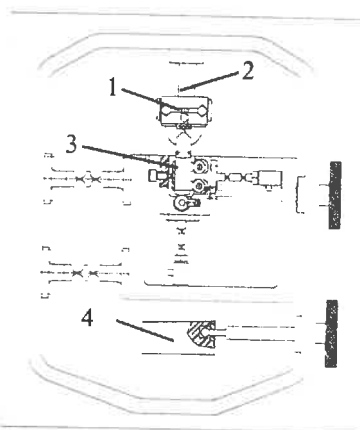


Figure 2 Assembly Drawing

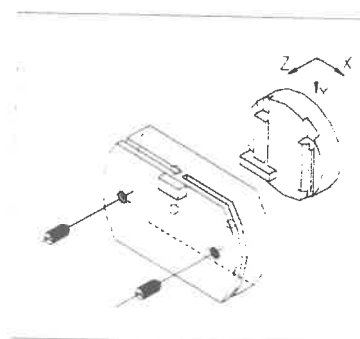


Figure 3 Sample Holder and Sample Mount

the purpose. The capacitance probe attached to the end of the PZT tube is used to measure the Z motions of the sample. The functional requirements of the PZT-capacitance probe assembly are, (1) to be thermal stable during operation, (2) fixing the target and setting the gap of $1\mu\text{m}$ for capacitance probe are to be easily achievable.

In order to ensure that the heavy electric field due to the high voltage applied to the PZT tube does not affect the capacitance probe, the capacitance probe was shielded using a thin macor tube. For thermal stability the cap probe is glued to the macor tube at only one end (at sample end) and the other end of the macor tube is glued to the outer disc. A 15mm x 15mm silicon wafer at the end of the sample mount serves as the cap probe target.

TOOL HOLDER AND FORCE SENSOR ASSEMBLY

The components of the tool holder and force sensor assembly are shown in figure 3. A single point diamond tool is brazed at the end of brass tool shank. The tool can be moved in the Z direction by rotating an eccentric cam with 0.5mm eccentricity. The tool can be moved in X direction using the two other indexing cams to make many cuts over the entire length of the sample. Two force sensors to measure cutting (X axis) and thrust (Z axis) forces are mounted to the tool assembly through circular flexures as shown. These double notch circular flexures made of beryllium copper are used to decouple the Z and X axis force measurement systems.

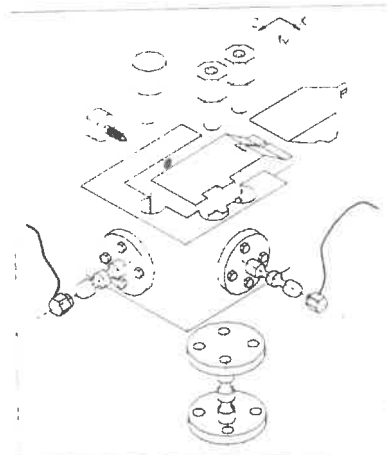


Figure 4 Tool Holder and Force Sensor Assembly

FINE POSITIONING SYSTEM

The function of fine positioning system is to provide a Z axis fine adjustment up to $50\mu\text{m}$ and Y axis fine positioning up to $500\mu\text{m}$ microns. These Z and Y axis fine adjustments are required during the tool-sample contact establishment procedure as described in the next section. Fine positioning in Z and Y axis is provided by means of wedges and four dual-axis flexures attached to the subframe. The flexures have a notch thickness of 1mm and are arranged on one side of the instrument to reduce the overall size (Refer to Figure 2).

THERMAL COMPENSATION

Since the structural loop is composed of different materials, thermal compensation of the loop is needed to keep the gap between the tool and sample constant. This was done by matching the thermal expansion of outer frame (Stainless steel grade 316) and the inner components. The material of inner components were carefully chosen and a block of aluminum was added so that the total expansion matched that of outer frame.

INSTRUMENT OPERATION

The actual cutting operation starts with establishing the tool-sample contact. For this, the tool is first brought close to the sample by using the eccentric cam provided in the tool holder assembly. Using this coarse positioning, the tool can be brought within 10 μ m with visual (light microscope) observation. Then the sample is vibrated at some suitable frequency and the tool is moved closer to the sample using the fine positioning system. The output of the thrust force sensor is monitored with a lock-in amplifier. Once the tool comes in contact with the sample, the force sensor output signal will correspond to the oscillation frequency of the sample. If the penetration depth is well within the elastic limit (say less than 1nm), no damage will be caused to the sample by this procedure. Once the tool-sample contact is established, cutting motions are provided to the sample by controlling the voltage applied to the PZT tube.

CONCLUSION

A second generation nanometric cutting instrument prototype was designed and fabricated. The main challenge for the design of the instrument was to comply with stiffness, overall size, thermal stability requirements and component material selection. The design has many new features and improvements over the first generation nanometric cutting instrument including better sample and tool access, better stiffness and thermal stability.

REFERENCES

1. Moriwaki, T, "Machinability of Copper in Ultra-Precision Micro Diamond Cutting", *Annals of CIRP* 38, 1, 1989.
2. Woirgard, J. and Dargenton, C. J., "A new proposal for design of high-accuracy nano-indenters" *Measurement Science and Technology*. 6 (1995), 16-21.
3. Von Turkovich, B. F., and Black, J. T, "Micro-machining of Copper and Aluminum Crystals", *Transactions of ASME, Journal of Engineering for Industry*, February 1970.
4. Lucca, D. A., and Seo, Y. W., "Effect of tool edge geometry on the resulting forces and energies in the orthogonal flycutting of TeCu", *Annals of CIRP* 42, 1, 1993.
5. Balasubramaniam, S., "Design of Nanometric Cutting Instrument", Masters thesis, University of North Carolina at Charlotte, 1995.

A FLEXIBLE ALIGNMENT FIXTURE FOR THE FABRICATION OF REPLICATION MANDRELS

James F. Cuttino

Dale E. Schinstock

Michael W. Todd

Department of Mechanical Engineering
The University of Alabama

ABSTRACT

A Flexible Alignment Fixture (FAF) has been developed to align a multi-piece replication mandrel for fabrication on a precision diamond turning machine (DTM). The fixture allows six degrees of motion of the workpiece relative to the spindle of the DTM. These degrees of freedom are necessary to remove error motions in the mandrel that are introduced during successive assembly/disassembly steps executed during fabrication of the mandrel. Requirements for small depths of cut necessitate that the axis of the mandrel be aligned with the axis of rotation of the chuck to within 25 μm error. Using conventional diamond turning methods, however, the minimum error motion obtained is often as high as 100 μm , rendering the mandrel useless.

The *FAF* replaces the vacuum chuck on the DTM and is used to precisely align the axis of the mandrel to the axis of the DTM. The fixture utilizes a hexapod arrangement to provide high stiffness, reduced error, and simple construction. The flexure has been used to realign a mandrel for the fabrication of an optic for NASA's X-ray Calibration Facility (XRCF), reducing the error motion of the mandrel from over 100 μm to less than 8 μm . Parts fabricated with the flexible alignment fixture show visible improvements in surface finish over parts cut on the original vacuum chuck.

INTRODUCTION

NASA uses precision diamond turning technology to fabricate replication mandrels for its X-ray Calibration Facility (XRCF) optics. As shown in Figure 1, the XRCF optics are tubular, and the internal surface contains a parabolic profile over the first section and a hyperbolic profile over the last. The optic is fabricated by depositing layers of gold and nickel on to the replication mandrel and then separating the resulting shell from the mandrel. Since the mandrel serves as a replication form, it must contain the inverse image of the surface. Figure 1 shows the components of the fabrication mandrel that are necessary to facilitate mounting, handling, and polishing during the fabrication process. More information concerning the replication mandrels has been documented by Fawcett and Engelhaupt [1993].

The difficulty in aligning the mandrel comes from the fabrication steps which it undergoes. The mandrel is rough machined and heat-treated prior to diamond turning. After diamond turning, silicon rubber separators that are undercut in radius by 3 mm (0.12") are inserted between the two end caps of the mandrel to allow the plating to wrap around the ends (to prevent flaking). The mandrel is then plated with a nickel-phosphor alloy using an electroless nickel process. At this point, the separators are removed and the mandrel is reassembled for the final cut on the DTM. The mandrel is measured for profile and finish, and polished to achieve an acceptable surface finish.

Wrapping the plating around the edges helps to prevent flaking, but it also destroys the alignment surfaces between the parts of the mandrel that insure that the axes of the parts are coincident. Several mandrels have been realigned by trial-and-error methods, consuming significant amounts of setup time. When the mandrel studied in this paper was reassembled, multiple efforts resulted in a minimum radial error motion of 100 μm . Since 50 μm of nickel plating was to be removed, and a minimum plating thickness of 25 μm was to remain on the part, the radial error motion had to be reduced to less than 25 μm . The mandrel was therefore unusable in its current state.

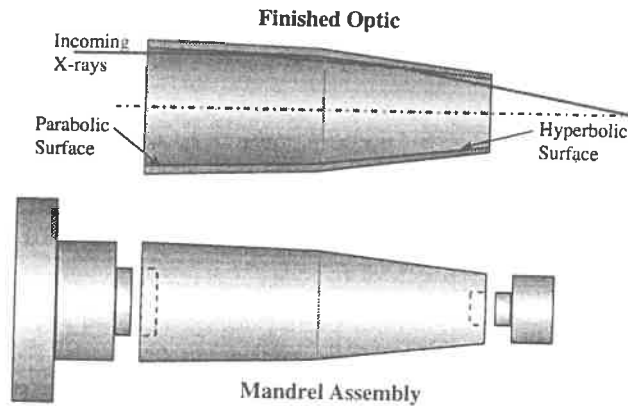


Figure 1: Geometry of XRCF Optical Surface and Replication Mandrel

DESIGN OF FIXTURE

To provide the motion necessary to align the mandrels, the *Flexible Alignment Fixture*, or *FAF*, was designed to replace the vacuum chuck normally used on the DTM. The design goals were to allow for linear displacements of $250\text{ }\mu\text{m}$ and rotations of 0.003 radians at the platform of the alignment fixture (over a 8" long mandrel, 0.003 radians results in an equivalent translation of one end of $600\text{ }\mu\text{m}$, or $0.024''$).

The design utilizes a six-legged hexapod arrangement to provide full flexibility and high stiffness. Figure 2 is a conceptual drawing of the *FAF* mounted on a diamond turning machine. The fixture has a base that mounts directly to the spindle bearing and a platform suspended by the six legs. The workpiece is mounted directly to the platform.

Since the required motion was very small, the flexure was designed to rely on deformation of the legs rather than joints to allow relative motion of each leg. The stress was calculated for each leg based on a maximum deflection, and the legs were designed to provide a factor of safety of 8 to 10. The legs were threaded on each end and screwed directly into the base and the platform, and preloading was achieved by lowering the platform towards the base, thus introducing a bending moment in each leg.

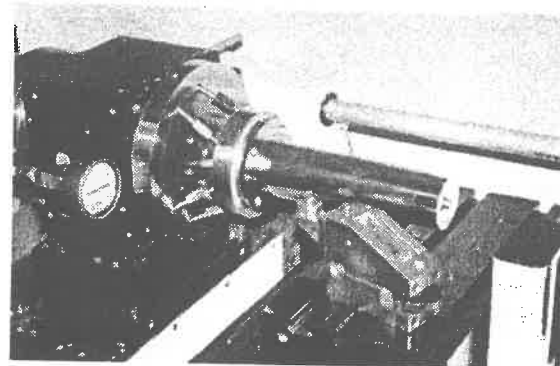
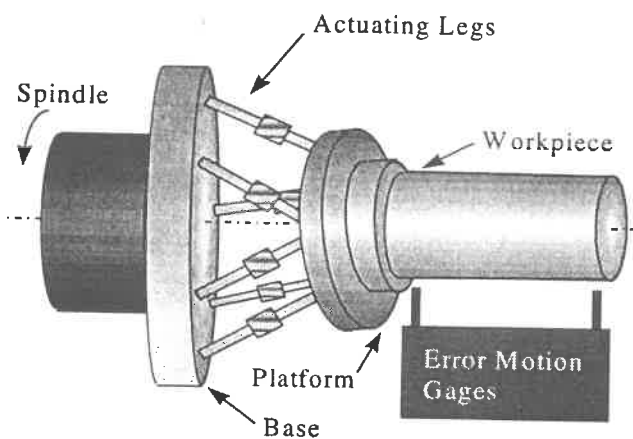


Figure 2: Flexible Fixture on a Lathe Spindle

The legs of the *FAF* utilized a differential pitch arrangement to provide high-resolution motion. A 3/8-24 UNF thread was machined on the portion of the leg inserted into the base of the *FAF*, and a 7/16-20 UNF thread was machined on the end inserted into the platform. This differential pitch in the screw provides a linear displacement of 210 μm per turn of the leg. In order to facilitate assembly of the *FAF*, the legs were fabricated in two sections joined in the middle by a hex nut. The legs were fabricated from AISI 410 stainless steel and were heat treated for additional strength.

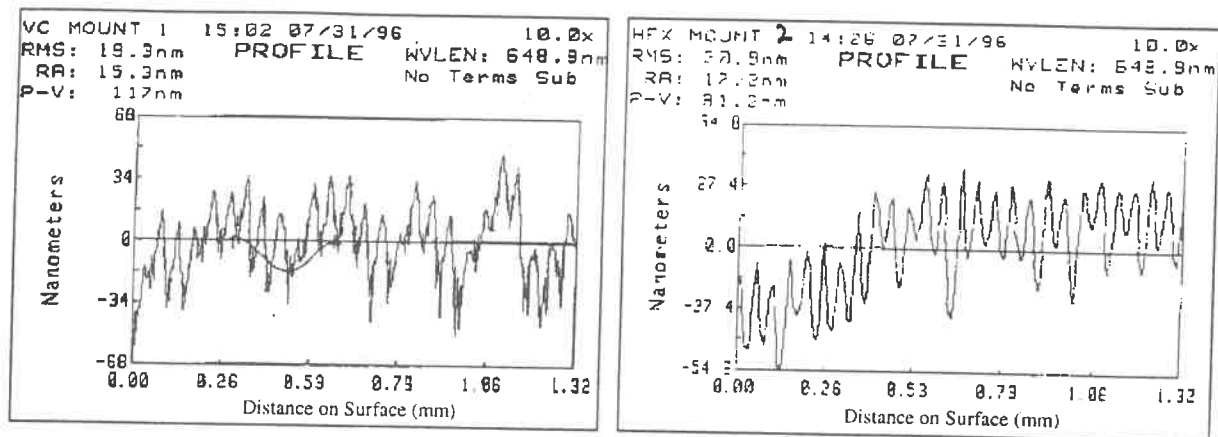
EXPERIMENTAL RESULTS

The vacuum chuck of the DTM was replaced with the *FAF* for testing. When mounted on the original equipment vacuum chuck, the test mandrel had a minimum error motion of 100 μm , which was only achieved after significant effort to align the mandrel pieces. Mounting the mandrel on the *FAF* provided the necessary freedom to align the axis of the mandrel to that of the DTM spindle. Once mounted, the error motion was measured by two dial indicators at four equally spaced rotations of the spindle to determine the proper adjustment. By placing the indicators at determined positions along the axis of the mandrel, two measurements at each of the four rotational positions provided the necessary error motion information. A kinematic analysis was conducted on the hexapod to provide the necessary corrections while minimizing leg adjustment so that stresses in the legs would be minimized. The resulting analysis was programmed into MATLAB to facilitate adjustments on the fixture. To adjust the *FAF*, the measurement values from the dial indicators were input into the kinematic program and the required leg rotations were output. With the *FAF*, the same mandrel was aligned to within 8 μm error motion with minimal effort, representing a 92% reduction in the error. The 8 μm error was significantly less than the target requirement of 25 μm .

There were several measures for success in the project. The most important was the ability to diamond turn the mandrel without penetrating the nickel plating. If the turning process penetrated the nickel plating, the mandrel would be scrapped. Other measures of success were the surface finishes and form errors (low frequency errors) that were achieved using the *FAF*. Although surface finish and form can be altered by post-polishing, this process is expensive and time consuming. A goal of the project was therefore to achieve surface finishes and form errors that were comparable to or better than those turned using the conventional DTM chuck. The mandrel cut on the *FAF* was therefore compared to one that had been successfully aligned on the vacuum chuck without the *FAF*. The two tests were conducted at a slide feedrate of 3 mm/min (0.05 mm/s) and a spindle speed of 1000 rpm (17 Hz).

Visual inspection of the surface finish on the mandrel cut on the *FAF* surprisingly showed significant improvement over the finish of the mandrel cut on the vacuum chuck. This conclusion was verified using a WYKO *Topo-2D/Topo-3D* surface interferometer. Figure 3 shows the surface profiles of the benchmark and *FAF* mandrels over a range of 1.32 mm. As shown in Figure 3a, the mandrel cut on the vacuum chuck has several components of error. The scan appears to consist of three fundamental components with wavelengths of approximately 0.25 mm, 0.05 mm, and a high frequency noise with wavelength of around 0.001 mm (measured along the axial direction). Based on the feedrate, the surface components had frequency components of 0.2 Hz, 1 Hz, and 50 Hz respectively. The rms amplitude of the error was 19.3 nm. A complex mounting configuration unfortunately prohibited more than one scan of this specimen.

Several traces were taken of the mandrel cut with the *FAF*. Figure 3b shows one trace which is representative of all of the traces taken. A significant discrepancy was realized in the amplitudes of the error, which ranged from an rms value of 17 nm to 60 nm. Since only one trace was made of the benchmark mandrel, the amplitude data is considered inconclusive. The trace shown in Figure 3b has an rms amplitude of 20.9 nm and shows an error component with a 1 Hz frequency (wavelength = 0.05 mm) superimposed on a very low frequency trend (possibly not even periodic) that continues outside the scan range of the interferometer. While the discrepancies in the results require further investigation, it appears that using the *FAF* actually reduced the 0.2 Hz and 50 Hz error motions seen in the benchmark mandrel. More data obtained from the mandrel cut on the conventional chuck may show similar form amplitudes, so the form error data (low frequency component) are not considered conclusive at this point. The elimination of the 50 Hz component in the mandrel cut on the *FAF* explains the visible improvement in surface finish on the part.



a) Mandrel cut on Vacuum Chuck

b) Mandrel Cut on *FAF*

Figure 3: Surface Profiles of Mandrels

Both parts are currently being measured using a *Long Trace Profilometer* to determine the actual error in each part. The preliminary results above, however, seem to indicate that the hexapod may help to reduce the error motion while salvaging the part.

CONCLUSION

A six-degree-of-freedom Flexible Alignment Fixture (*FAF*) has been developed to adjust the position and orientation of a workpiece on a diamond turning machine spindle. The device uses a hexapod design and kinematic analysis to reposition the workpiece. The designed range of motion allows for displacements of 250 μm and changes in orientation of 0.003 radians.

A fabrication mandrel for NASA's XRCF project, which was unusable due to alignment problems, has been salvaged using the *FAF*, which was able to reduce the error motion from 100 μm to 8 μm . The *FAF* also greatly reduces the effort necessary in insuring the alignment of the mandrel components prior to cutting. Preliminary results indicate that improved surface finish may be an added benefit to using the *FAF*. Pending results from a long trace profilometer will yield more information on the effectiveness of the *FAF*.

ACKNOWLEDGEMENTS

Sincere thanks go to NASA's Marshall Space Flight Center and to Dr. David L. Lehner, Mr. William D. Jones, Mr. John Cernosek, and Mr. Kent Bachelor for their input and support on this project.

REFERENCES

Fawcett, S.C., and D. Engelhaupt, 1993, "Production of X-Ray Optics by Diamond Turning and Replication Techniques," *Proceedings of the American Society for Precision Engineering*, Volume 8, pp. 95-98.

LOW-COST METROLOGY OF 6-DEGREE-OF-FREEDOM, MILLIMETER-RANGE STAGES

Yan Wu and James F. Cuttino
Department of Mechanical Engineering
The University of Alabama

This abstract describes a new low-cost metrology system for 6-degree-of-freedom millimeter-range stages. The objective of this measurement system is to achieve metrology resolution in the sub-micrometer regime. Toward this end, the performance characteristics of the sensors used in the metrology system are analyzed and evaluated experimentally.

1. Introduction

Microscanning mechanisms are becoming more common for a variety of applications including STM imaging¹, optical lithography with step-and-repeat cameras², and scanning test stands for magnetic read/write head research³. Conventional means for detecting the three-dimensional (3-D) behavior of a scanning stage typically use capacitance gages and/or laser interferometers. However, these instruments can be very complex, expensive and difficult to set up. Simpler and less expensive metrology systems are not only desirable for research, but also beneficial for industrial applications

This article describes the development of a low-cost method for conducting non-contacting, three-dimensional metrology in the sub-micrometer regime. The system uses three four-quadrant photo detectors to detect the positions of incident beams from three laser diodes attached to the measurement target. The system is designed to measure position and orientation of a multi-degree-of-freedom stage within ± 1 mm travel. The total cost of the system is less than \$1000.

In our previous work⁴, we discussed the detailed kinematic analysis of the system including working envelope and error budget prediction. In this paper, we present the theoretical analyses and experimental results of performance characteristics of the laser diode / photo detector pair.

2. Description of Metrology System

The metrology concept is illustrated in Figure 1. The system uses 3 photo detectors mounted in specified orientations as position-sensitive detectors. Each of the photo detectors has four quadrants that measure the intensity of light landing on that particular quadrant. Three laser diodes are mounted to the object being measured and project their beams onto the photo detectors. As the body being measured moves in space, the corresponding positions of the incident beams are measured by the photo detectors. Using kinematic

relationships developed at The University of Alabama, position and orientation information of the object is calculated from the output currents of the photo detectors.

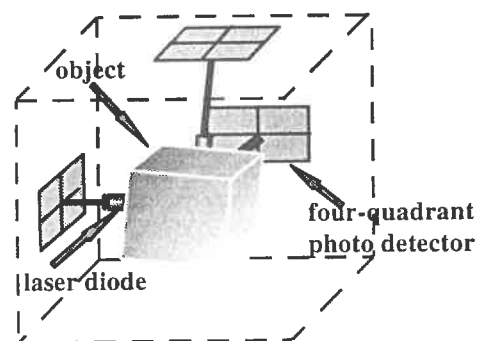


Figure 1: System Configuration

The error budget presented by Prather, Cuttino, and Schinstock⁴ shows the dependence of the system performance on the ability to accurately locate the incident laser beam on the photo detector. This paper describes work being conducted to characterize the diode/detector performance.

3. Sensor performance evaluation

A four quadrant photo detector was selected to be the position sensor because of its high resolution and low-cost⁵. Assuming the light spot to be circular and of uniform intensity, the position of the beam can be calculated from the measured output currents of each of the photo-detectors using the following relationships:

$$\sin^{-1}\left(\frac{2S_x}{d}\right) + \frac{2S_x}{d} \sqrt{1 - \left(\frac{2S_x}{d}\right)^2} = \frac{\pi (i_1 + i_4) - (i_2 + i_3)}{2 (i_1 + i_2 + i_3 + i_4)} \quad (a)$$

and

$$\sin^{-1}\left(\frac{2S_y}{d}\right) + \frac{2S_y}{d} \sqrt{1 - \left(\frac{2S_y}{d}\right)^2} = \frac{\pi (i_1 + i_2) - (i_3 + i_4)}{2 (i_1 + i_2 + i_3 + i_4)} \quad (b)$$

As depicted in Figure 2, S_x and S_y are horizontal and vertical deviations of the light spot from the center of the detector. i_1, i_2, i_3 and i_4 are the average currents of quadrants I, II, III, IV, respectively. d is the diameter of the light spot.

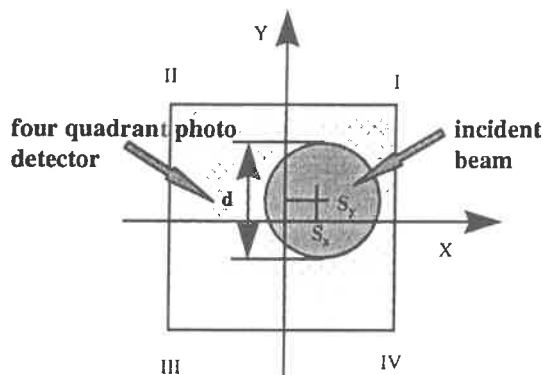


Figure 2: Geometric Definition for Beam on Photo Detector

Equations (a) and (b) can be expanded in series to yield solutions for S_x and S_y . These expressions are given by

$$S_x = \frac{d}{2} \left(\frac{Z_x}{2} + \frac{Z_x^3}{48} + \frac{13Z_x^5}{3840} + O(Z_x^6) \right) \quad (c)$$

and

$$S_y = \frac{d}{2} \left(\frac{Z_y}{2} + \frac{Z_y^3}{48} + \frac{13Z_y^5}{3840} + O(Z_y^6) \right) \quad (d)$$

where Z_x and Z_y are

$$Z_x = \frac{\pi (i_1 + i_4) - (i_2 + i_3)}{2 (i_1 + i_2 + i_3 + i_4)} \quad \text{and} \quad Z_y = \frac{\pi (i_1 + i_2) - (i_3 + i_4)}{2 (i_1 + i_2 + i_3 + i_4)}$$

For some applications, the first order solution is accurate enough⁶. We found that a fifth order solution reduced errors to less than $1\mu\text{m}$ for $\pm 0.78\text{ mm}$ travel, but a higher order fit is required to improve the resolution to sub-micron levels.

Equations (a) and (b) are based on several assumptions. First, the light spot must be within the range of the active surface of the photo detector so that none of the beam falls outside of the detector. Second, the shape of the light spot is assumed to be circular. This is not valid for most laser diodes or when the incident beam is not perpendicular to the photo detector. Finally, the intensity distribution is assumed to be uniform, although the distribution is actually Gaussian. All of these assumptions are directly related to the performance characteristics of the system.

4. Experimental Procedure

A test stand has been built to characterize the performance of the photo detector and laser diode pairs. Shown in Figure 3, the test stand consists of a Dover X-Y air bearing stage

controlled by a personal computer. A Zygo AXIOM 2/20 laser measurement system is used to give the actual position of the stage. The laser diode is fixed above the stage, and the photo detector is mounted on the stage. By moving the stage, the output from the photo detector can be mapped and the stability of the laser diode characterized.

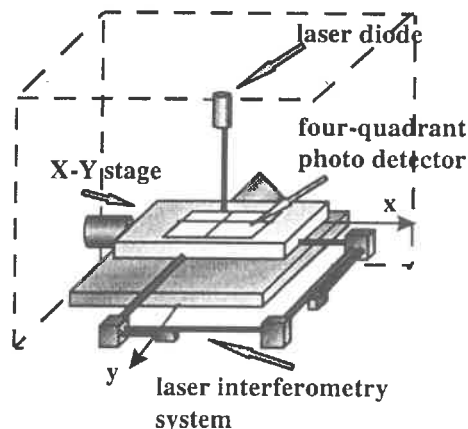


Figure 3: Laser diode /Detector Mapping

As illustrated in Figure 4, three experiments were conducted to test the performance of the position sensor. The first experiment measured sensor stability. Experiment 2 tested the ability of the diode/detector pair to measure position. Experiment 3 measured the effect of a non-uniform, asymmetric laser beam on the system. This non-symmetric beam is shown in Figure 4.

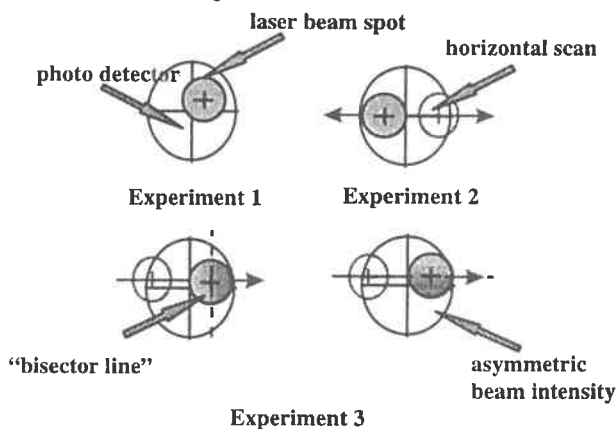


Figure 4: Experiment Design

The purpose of the first experiment was to test the stability of the sensor signal by keeping the laser beam at a fixed position on the photo detector. The beam diameter d is 3.2 mm . The data sampling rate is 210 Hz . Figure 5 shows that the error signal is $10\mu\text{m}$ over of period of 6 seconds. The magnitude of the error signal was reduced to half in later experiments after the test stand was isolated from ambient light source. This result shows the need for either optical

filtering or signal processing to achieve sub-micron resolution.

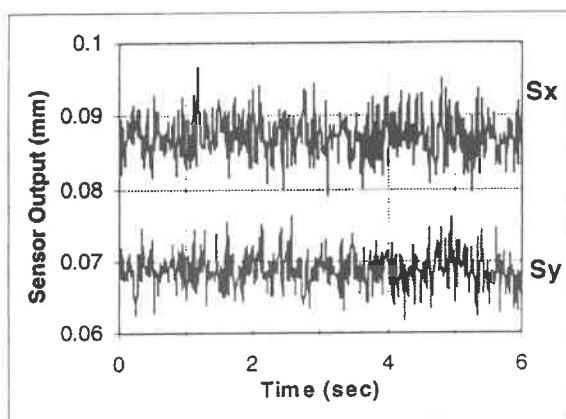


Figure 5: Result from Experiment 1 Showing Sensor Output Stability

In the second experiment, the stage was moved forward and backward along the X direction in order to let the laser beam scan across the active area of photo detector. The sensor output was compared with the output from the laser interferometer which has a measurement resolution of 1.25 nm. The laser diode was fixed throughout the experiment.

In Figure 6, the outputs from the photo detector and the laser interferometer are compared. Curve X is the laser interferometer output. Curve S_x represents the calculation result using equation (c) based on the measurements of the photo detector output currents. The curves show excellent agreement within the noise on the detector. To better compare the two signals, the calculated sensor output was plotted as a function of the stage position in Figure 7. Within the range of ± 1.3 mm, the curve is a nominally straight line with unity slope. The mapping is very repeatable as seen by the multiple passes in Figure 7. This shows that errors can be removed from the system through calibration. Experiment 2 shows that equations (c) and (d) are valid for the positioning device and that the output is repeatable.

In experiment three, the effect of the intensity distribution on the sensor output was tested. It is common that the intensity distribution of commercial laser diodes is not central symmetric. The laser diode used in this experiment was visibly asymmetric, with one half darker than the other and the "lighter" and "darker" sides bisected with an imaginary line through the center of the beam. To measure the effect on this system, we first put the "bisector line" perpendicular to the scanning direction. Then we put the "bisector line" parallel to the scanning direction. This is shown in Figure 4.

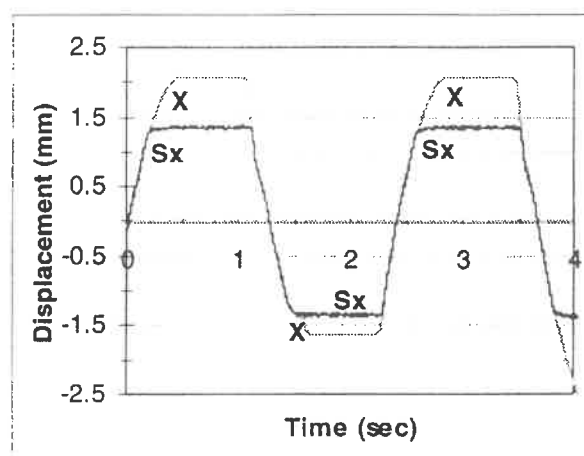


Figure 6: Result from Experiment 2 Showing Horizontal Scan Across Photo Detector

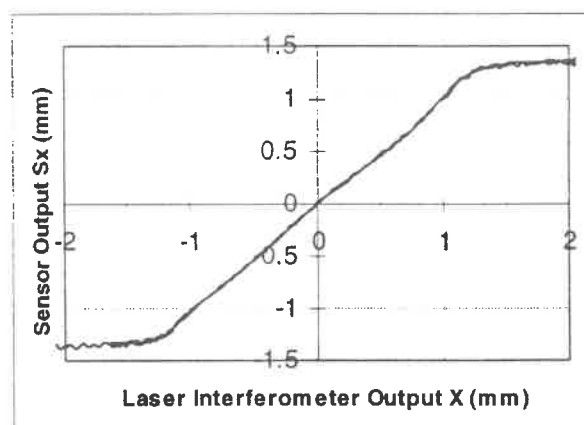


Figure 7: Sensor Output as a Function of the Stage Position in Experiment 2

The experiment results are shown in Figure 8 and Figure 9 respectively. These figures show that the intensity distribution of the incident laser beam has a great effect on the performance of the system. However, a linear relationship between the sensor output and actual displacement is still realized as long as the intensity distribution is symmetric along the scanning direction. This can be easily proven mathematically. A further implication is that the incident beam does not need to be uniform in intensity, but it is essential that the beam intensity is symmetric about the center in order to get two dimensional measurement from the photo detector and laser diode pairs. In this case, a calibration process before measurement is highly recommended since equation (a) and (b) is no longer valid theoretically. This form of calibration is routine in the AFM (atomic force microscopy) industry⁷ and should be applicable in this case as well.

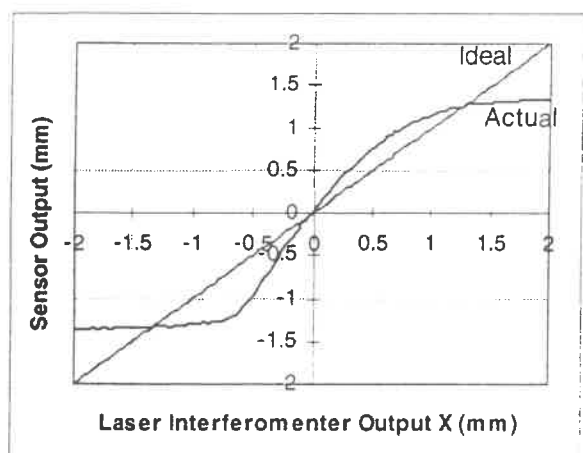


Figure 8: Experiment 3 Results when Bisector Line is Perpendicular to Travel

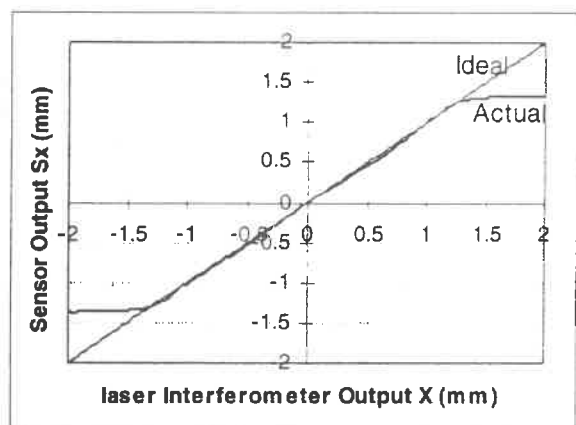


Figure 9: Experiment 3 Results when Bisector Line Parallel to Travel

5. Conclusion

The proposed metrology system will provide cost-effective, three-dimensional, non-contacting measurement in millimeter-range motion. The measurement quality of the system depends on the performance of the four-quadrant photo detector and laser diode pairs which are used as position sensors in this system. To evaluate the performance of the sensors, a test stand has been built and a series of experiments have been conducted at the Metrology and Precision Engineering Laboratory (MAPEL) at the University of Alabama.

Experimental results show that the proposed metrology system is not only feasible, but it also has great promise for achieving high resolution. The objective of future work is to optimize the performance of the position sensor by trying different kinds of laser diode and photo detector pairs.

References

- ¹ Holmes, M., R. Hocken, and D. Trumper, 1996, "A Long - Range Scanning Stage Design," *Proceedings of The American Society for Precision Engineering*, Vol. 14, pp. 322-327.
- ² Nakamura, Osamu and Nitsuo Goto, 1994, "Four-beam Laser Interferometry for Three-dimensional Microscopic Coordinate Measurement", *Applied Optics*, Vol. 33, No.1, pp. 31-36
- ³ Schults, S., "Five Degree of Freedom Scanning Test Stand for Magnetic Read/Write Head with 1 nm Resolution," *Proceedings of the ASPE Spring Topical Meeting*, Vol. 11, April 1995
- ⁴ Prather, M. J., J. F. Cuttino, and D. E. Schinstock, "Three-dimensional Metrology Frame For Precision Applications," *Proceedings of the 1996 American Society for Precision Engineering Conference*, Vol.14, pp. 668-671.
- ⁵ Constant A. J. Putman et al., "A detailed analysis of the optical beam deflection technique for use in atomic force microscopy," *Applied Physics*, Vol. 72, July, 1992, pp. 6-12.
- ⁶ Anssi J. Mäkynen, Juha T. Kostanovaara, and Risto A. Myllyä, 1994, "Tracking Laser Radar for 3-D shape Measurements of Large Industrial Objects Based on Time-of-Flight Laser Range Finding and Position-sensitive Detection Techniques," *IEEE Transactions on Instrumentation and Measurement*, Vol.43, No.1 Feb. 1994, pp. 40-49.
- ⁷ Neill P. D'Costa and Jan H. Hoh, 1995, "Calibration of optical lever sensitivity for atomic force microscopy," *Review of Scientific Instruments*, Vol. 66, Oct. 1995, pp. 5096-5097

DEVELOPMENT OF A LONG RANGE, TRACTION DRIVE FAST TOOL SERVO FOR DIAMOND TURNING APPLICATIONS

*Michael W. Todd and James F. Cuttino
Department of Mechanical Engineering
The University of Alabama*

Introduction

Fast Tool Servos (FTS's) have allowed Single Point Diamond Turning Machines (DTMs) to become more versatile in the range of applications for which they are used. One major limitation of the FTS, however, is that the range of motion is typically less than 400 micrometers due to limitations in actuator range and precision. The goal of this work is to extend the range of motion of FTS's to enable the fabrication of more complex optics with deep, non-axisymmetric features. Target characteristics are a range of motion of one millimeter at 50 hertz.

Selection of Traction Drive Configuration

The main objective of this research is to increase the range of motion of the FTS. Several potential designs were considered which could meet this objective including a voice coil approach, hydraulic arrangement, and piezoelectric (PZT) designs. Along with range and frequency of motion, other factors considered were stiffness, manufacturability, and controllability. The traction drive configuration was selected because of virtually unlimited range and the mechanical advantage gained through the pinion radius. Traction drives represent a compromise between the high stiffness but limited range of PZT actuators and the low stiffness, larger displacement voice coil actuators. Traction drives in larger systems have been proven capable of providing submicron position accuracy [1,2].

Figure 1 displays the actuation scheme used in the work. The torque from the rotary actuator is transferred to a linear force on the tool shuttle through a steel ribbon that is wrapped around the shaft and attached to the tool shuttle. This design minimizes excessive normal loads and slipping in a typical friction drive system. It also simplifies the suspension design since the shuttle does not need to resist normal forces from the driver roller. The preload between the ribbon and the shaft is adjusted by increasing or decreasing the tension in the ribbon.

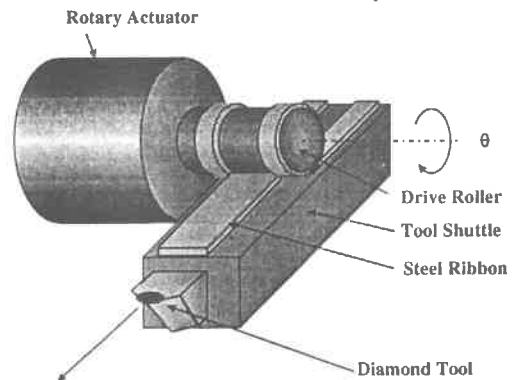


Figure 1: Traction Drive Servo

Suspension of the Tool Shuttle

A principal challenge in the long-range FTS development is designing a bearing system that provides the necessary motion with sufficient stiffness and accuracy. Conventional cantilever

flexure bearings could not be used because they would have to be excessively large to provide the desired range of motion while maintaining stresses low enough to avoid fatigue. Conventional annulus flexure bearings are too restrictive on the axial motion as well. Therefore, a new type of flexure design, shown in Figure 2 and Figure 3, was developed. The rolling action of the flexures allows improve compliance in the axial direction, and the flexures support loads in the transverse directions through shear as shown in Figure 3.

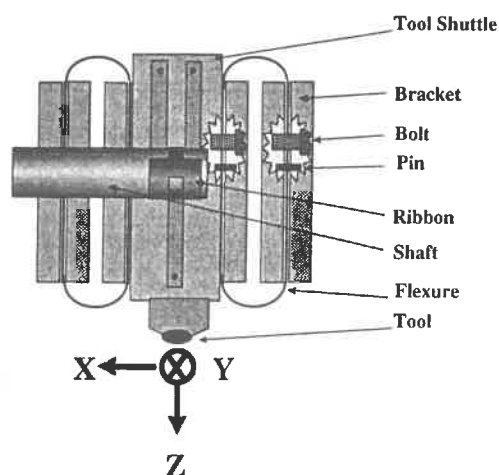


Figure 2: Flexure Design

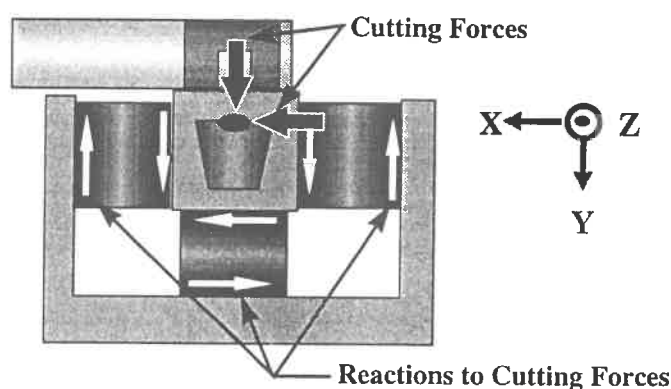


Figure 3: Forces on Flexures

Complex geometry and loading conditions of the flexures prevented closed form analysis of the flexures, so an empirical study was conducted using Finite Element Analysis (FEA). The FEA analysis was performed using ALGOR software. Only one flexure was modeled for each case due to symmetry. The loads on the flexure were varied and the displacements were obtained to prove that stiffness was linear. Then each of the flexure dimensions was changed to see its effect on flexure stiffness. The stresses were also computed using an input displacement.

The FEA analysis allowed development of various trends that helped in the design of curved flexures. Figure 4 shows that force is linear with respect to displacement for loading in the X, Y, and Z directions, so curved flexures will not produce any significant nonlinearities. Also, the Y-direction still maintains the highest stiffness, while the Z-direction is much more compliant. The effects of the flexure dimensions on stiffness were also examined. For all of the following trends one dimension was changed while the others were held constant. Parameters that were varied were flexure thickness, flexure height, and radius of curvature. Figure 5 shows the flexure dimensions of interest. Figure 6 shows that stiffness increases with respect to flexure thickness in a cubic trend. Figure 7 shows that increasing flexure height increases stiffness. This effect is

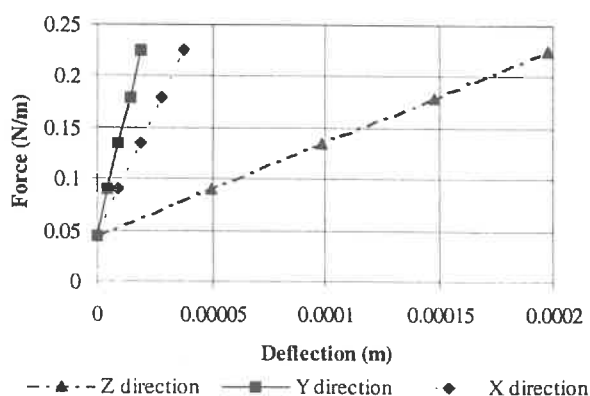


Figure 4: Force vs. Deflection

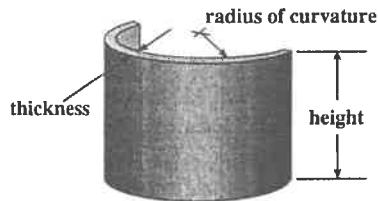


Figure 5: Flexure Dimensions

changes in dimensions which is desirable since the stiffness in that direction should come from the actuator rather than the flexures. Knowledge about these trends and information concerning the tool forces encountered in diamond turning [3] can be used to design curved flexures that are well suited for long range FTS applications.

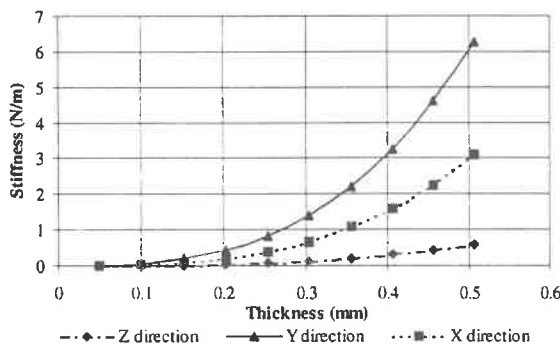


Figure 6: Effect of Thickness on Flexure Stiffness

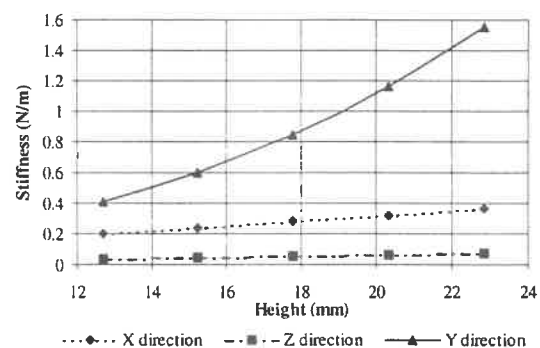


Figure 7: Effect of Flexure Height on Flexure Stiffness

Trends were also developed for stress so that the flexures could be designed to avoid plastic deformation or fatigue failure. Stress is linear with respect to displacement for X-, Y-, and Z-directions. For a given displacement, the lowest stress occurs for the smallest possible thickness and the shortest radius of curvature which still give the desired stiffness. The principal advantage of the curved flexures is that they can be made more compact than flat flexures for the same loads and displacements.

Prototype Device

The prototype device shown in Figure 8 was constructed for testing. A PID control algorithm was used to control the device. The computer sent a voltage out through the 12 bit D/A on a Computer Boards CIO DAS1602/16 card that was amplified using a Copley Controls Model 200 motor amplifier. The output of the amplifier was connected to a BEI model RA16-06 rotary actuator, and the position was read using an LVDT. The position was fed back to the computer through the 16 bit A/D on the Computer Boards card. A block diagram of the FTS control system is shown in Figure 9. The mass of the tool shuttle was 150 grams. The axial stiffness of the flexures was 3500 N/m, and the maximum torque of the rotary actuator was 0.07768 N-m.

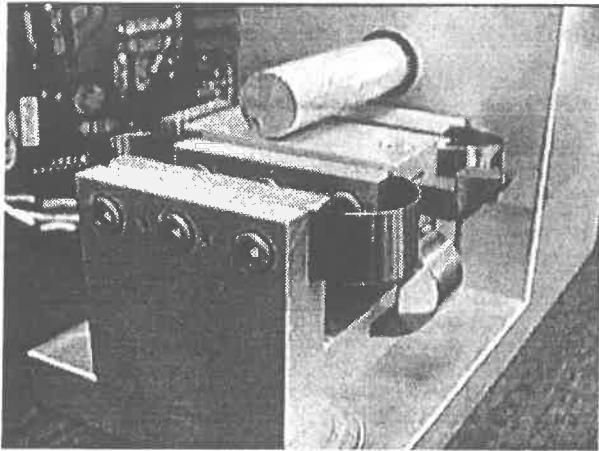


Figure 8: Prototype Traction Drive FTS

The system behaved as an underdamped second order system with a natural frequency of 21 hertz. The system easily generated a one millimeter amplitude up to 20 hertz. After the natural frequency, there was some attenuation and increased phase lag as expected. Work is currently being conducted to optimize PID gains for minimum error and maximum disturbance rejection. The preliminary results indicate that the traction drive will be capable of the desired range and frequencies.

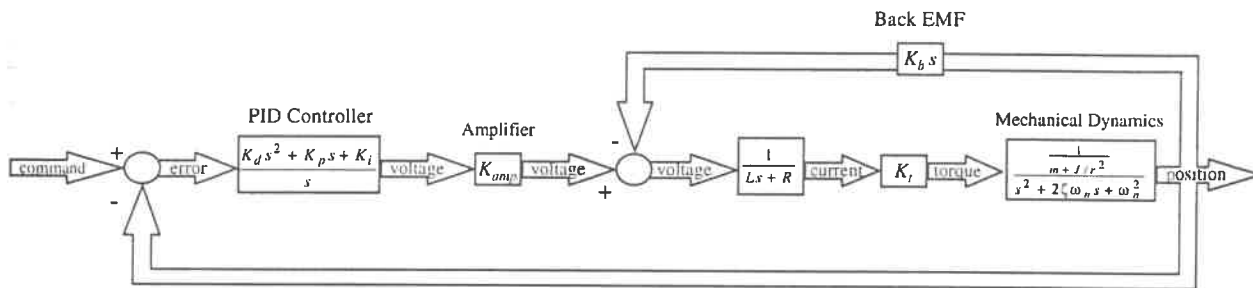


Figure 9: Block Diagram of the Prototype FTS System.

Conclusions

A prototype FTS based on a traction drive configuration has been constructed at The University of Alabama. Tests done on the device indicate that the traction drive is feasible for long-range FTS applications. A new flexure configuration has also been investigated using finite element analysis and experimental verification. Results from the study indicate that extended linear motion can be obtained from the new flexure design without significantly reducing stiffness. A second prototype device is currently being constructed using information that has been learned from the work conducted to date.

Acknowledgments

Special thanks go to NASA's Marshall Space Flight Center for support of this project.

References

1. Slocum, A. H., *Precision Machine Design*, Englewood Cliffs, NJ: Prentice-Hall, 1992.
2. Rao, G. S. and P. I. Ro, "Submicrometer control of a traction drive using state feedback and estimation," *Precision Engineering*, 17, (2), April 1995, 124-130.
3. Drescher, J. D. and T. A. Dow, "An Empirical Tool Force Model for Precision Machining," *Proceedings from ASPE 1996 Annual Meeting*, November 1996, 500-505.

FAUST AND WAGNER: MACHINE DESIGN AND PROCESS SIMULATION

O.W. Fähnle⁺, H. van Brug[#], C.J. van der Laan[#] and H.J. Frankena[#]

⁺ TNO Institute of Applied Physics, P.O.Box 155, 2600 AD Delft, The Netherlands

[#] Delft University of Technology, Dept. of Applied Physics, Optics Research Group,
Lorentzweg 1, 2628 CJ Delft, The Netherlands

ABSTRACT

Based on the same functioning principle, two new fabrication techniques, FAUST (Fabrication of Aspherical Ultraprecise Surfaces using a Tube; Patent pending) and WAGNER (Wear-based Aspherics Generator using a Novel Elliptical Rotator, Patent pending) have been developed. Their functioning principles have been described in detail in previous papers¹⁻³. Both combine the self-correcting character of traditional load-controlled grinding and polishing with features of high precision machine tools, featuring an advantageous error propagation factor between tool shape and generated surface shape. Shaping and finishing are carried out on the same machine, improving the eventual shape accuracy. The machine designs of two tools based on FAUST and WAGNER are presented and first tests with the tool based on WAGNER are discussed. Finally, an accuracy analysis is presented with which the influences of tool misalignments can be simulated together with a process simulation of the production process. The obtained results can be used for compensation.

1. FUNCTIONING PRINCIPLE

Let a radial section (r in the range of $\Delta \leq r \leq r_m$) of a rotationally symmetric convex or concave surface of revolution $f(r)$ be intersected by a plane that is parallel to the y , axis and that contains the points $P_\Delta = \{\Delta, 0, f(\Delta)\}$ and $P_m = \{r_m, 0, f(r_m)\}$ (the function $f(r)$ describes the cross-section of the surface in a plane which contains its axis of symmetry) [Fig.1]. The intersection curve between the surface and this plane is located in the x, y plane and will in the following be referred to as the *Shape*. Positioning this *Shape* above a rotating workpiece at an off-axis distance Δ and with a fixed angle α , given by

$$\alpha = \arctan\left(\frac{f(r_m) - f(\Delta)}{r_m - \Delta}\right), \quad (1)$$

with respect to the r axis, it can be used for production. Therefore, for the generation of a radial section $\Delta \leq r \leq r_m$ of the surface, the shape of the intersection curve (the *Path*) is given by¹

$$x(r) = \frac{f(r) - f(\Delta)}{\sin \alpha} \quad \text{and} \quad y(r) = \sqrt{r^2 - \left(\frac{f(r)}{\tan \alpha} + \frac{\Delta f(r_m) - r_m f(\Delta)}{f(r_m) - f(\Delta)}\right)^2}. \quad (2)$$

Consequently, it is possible to generate all kinds of concave and convex, on- and off-axis surfaces of revolution. The shape of the generated surface depends on the shape of the intersection curve as well as on the angle between tool and surface. It is hence possible by varying the angle α to use the same *Path* for the generation of different kinds of surfaces. While FAUST uses a line-contact between tool and workpiece along the *Shape*, WAGNER is applying a point-contact. Both, FAUST and WAGNER apply load-controlled grinding and polishing.

2 FAUST

A tool based on FAUST is currently being constructed. Within the stiff geometrical frame of a milling machine a state-of-the-art air bearing spindle (with an absolute rotational error of 50 nm) is used to rotate the workpiece (e.g. BK7) around a vertical axis. A tubular element T [Fig.2] that has a non-circular cross-section in a plane that is parallel to the x, y plane is fixed above the workpiece in such a way that the f and the z axis intersect and form the required angle α . The surface is processed

using a machining band that is guided along the lower rim of the tubular element applying a line-contact with the workpiece in the x, y plane. A circular element D that is rotating around an axis parallel to the z axis is driving the band and pressing it radially outwards by the use of an elastic element R (e.g. a rubber ring). That way the machining band is following the non-circular cross-section of the tubular element and tool and workpiece are in line-contact along the required *Shape*.

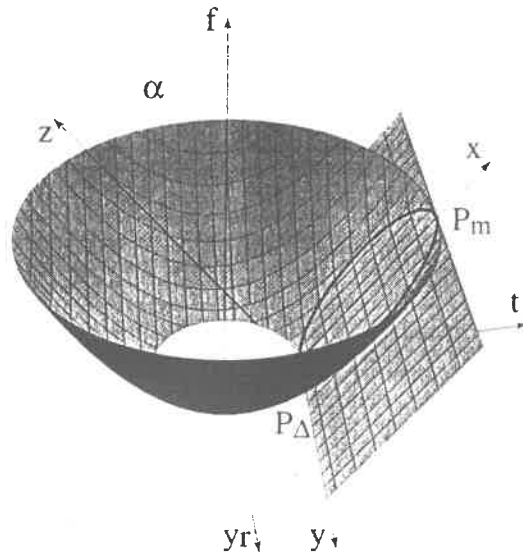


Fig.1 Functioning principle of FAUST and WAGNER and definition of the y_r, t, f system of coordinates of the workpiece and the y, x, z system of coordinates of the *Shape*.

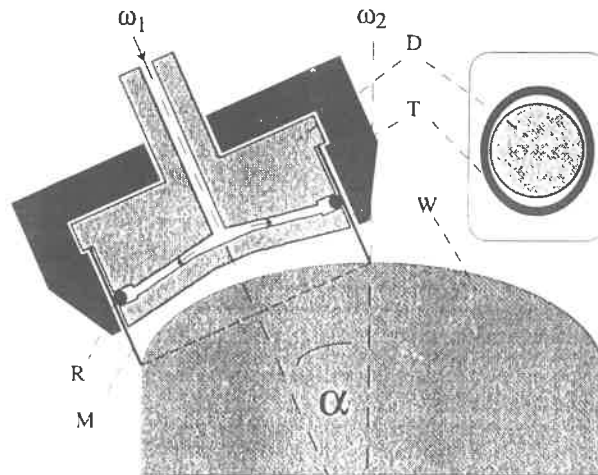


Fig.2 FAUST tool design: T indicates the fixed tubular element, M the machining band, W the workpiece and D a rotating circular element that is driving the band and pressing it by the use of the elastic element R (e.g. a rubber ring) radially outwards against T . (the inset shows a top view)

The tool has one translational degree of freedom such that the *Shape* moves parallel to the f axis down into the material. In order to achieve this the loaded tubular element has a translational degree of freedom along an axis that is located in the x, z plane and that has an angle $\alpha - \sigma$ with respect to the f axis (the angle σ depends on α and on the ratio between tool wear and workpiece wear¹). As FAUST applies a load controlled process, it enables loose abrasive ductile grinding and subsequent bowl-feed polishing to generate aspherical surfaces.

3 WAGNER

A small tool is mechanically guided by the use of a tool-path shaping element and is applying a point-contact along the *Path*. As the machining process is load controlled and the relative speed between the loaded tool and the workpiece varies along the *Path*, the wear is not constant along the tool path. Therefore, a mechanical stop is applied to ensure that the process stops in the x, y plane where the *Shape* is located (at $z = 0$). Based on WAGNER, a tool was built capable of producing all kinds of on- and off-axis conic surfaces (surfaces of revolution with conic sections as generators)^{2,3}. The tool-path shaping element is based on a geometrical effect using three pins of the same radius r_o [Fig.3]: if a pin P is rotated around the two fixed pins F_1 and F_2 and at the same time is pressed radially outwards by the use of a spring against a belt of constant length l that is encircling all three pins, the center of P follows a geometrically perfect ellipse in the x, y plane. In this way, an elliptical *Path* is generated that is determined by the distance e between F_1 and F_2 and by the inner length l of the belt. The center of this ellipse is positioned at the rotary axis R of the tool-path shaping element and the three pins are located parallel to the z axis. The rotary axis R must have a diameter of $\delta < 2r_o$ in order to avoid contact with the belt. Adjusting the line ε through the two points F_1 and F_2 , either parallel to the y axis or parallel on the x axis all kinds of off- or on- axis conic surfaces can be generated³ not requiring an in-process tool-path control. In order to enable the production of different scales of the same surface, a pantograph is used that is mechanically connecting the points

P , R and S . The pantograph is rotating in a plane parallel to the x, y plane around its fixed reference point R . Its stiff construction [Fig.3] allows the use of the pantograph to drive the pin P as well as to guide S along a scaled version of the path of P . The arm ξ is fixed to R and is rotating with a constant angular speed. Therefore, the points P and S are accelerated along their tool-paths, thereby avoiding the generation of cycloids. The machining is carried out at S using fixed abrasive load controlled point-contact grinding and polishing using a small cylindrical machining-head that is rotating around its axis of symmetry. The machining head can translate along an axis parallel to the z axis. It is pressed against the workpiece surface by the use of a spring and is machining the surface until it reaches the mechanical stop. The machining head is then rotating along the required *Shape* in the x, y plane. The *Shape* accuracy is thus given by the rigidity of F_1 and F_2 with respect to the workpiece coordinate system, by the accuracy of the belt length (which depends on its longitudinal elastic constant), by the accuracy of r_a and by the accuracy of the scaling factor of the pantograph. A positioning inaccuracy of the rotary axis R with respect to the center of the generated ellipse does not affect the production. As relatively slow rotational speeds ($\omega_1 \approx \omega_2 \approx 1\text{ Hz}$) and small forces (P is pressed with about 20 N outwards against the belt and the machining head is pressed with about 10 N against the workpiece surface) are applied and the tool will be used within the stiff geometrical frame together with a state of the art air bearing spindle for the workpiece rotation, highest accuracy levels can be expected.

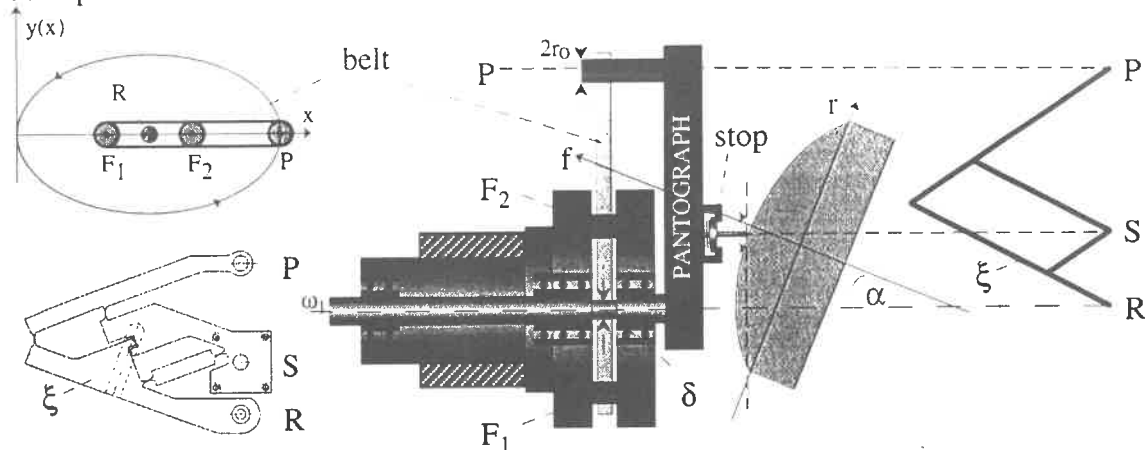


Fig.3 Sketch of a tool based on WAGNER for the generation of conic surfaces. The pantograph consists of a 15 mm thick aluminium plate comprising 4 pivot holes that are generated by electric discharge machining enabling a play-free movement of the pantograph.

In order to demonstrate the possibility to generate a smooth surface applying WAGNER, a first test was carried out implementing the tool on a not state-of-the-art milling machine that was poorly isolated from vibrations and other mechanical disturbances and that comprised a conventional workpiece spindle. The amplitude of the machine vibrations that causes distortions to the *Shape* was measured to be 5-10 μm at 1 Hz corresponding to the rotary speed of the tool. The surface (glass, BK7, with $r_m=70\text{ mm}$) was microground using cup wheels having a diameter $d = 10\text{ mm}$, using 1 μm loose abrasive diamonds, an oil-based slurry and tin as cup wheel material. Once a smooth surface was achieved, the process was stopped and it was measured to have an average roughness of $R_a = 4.15\text{ nm}$. The fitted conic surface, a hyperboloidal surface, is within the adjustment accuracy of the angle α . The shape inaccuracy of the fitted conic surface of 9.7 $\mu\text{m PV}$ (standard deviation $\sigma = 4.89\text{ }\mu\text{m}$) with respect to the measured surface can be explained by the amplitude of the machine vibrations. Consequently, future experiments will be carried out on a better grinding machine using a more appropriate type of belt and better allignment possibilities.

4 DISCUSSION

Likewise a curve generator that uses a cup-wheel to generate spherical surfaces of revolution, FAUST and WAGNER are characterized by their three degrees of freedom: the rotation of the workpiece

around its symmetry axis, the rotation of the tool along the *Shape* and the translation of the *Shape* parallel to the f axis down into the material. Both FAUST and WAGNER are generating the *Path* mechanically in advance not requiring an in-process control of a complicated tool path. It is not necessary to achieve the same shape accuracy for the *Shape* as is required for the workpiece: there is an advantageous error propagation factor between the shape of the intersection curve and the workpiece that depends on the angle α^1 . The shape accuracy of the generated surface depends on the shape accuracy of the *Shape* as well as on its accurate position and orientation with respect to the workpiece. Therefore, an isolation from vibrations and temperature changes is required. On the other hand, a tool based on WAGNER or FAUST is less sensitive to external mechanical disturbances than a precision grinding machine, because the tool is held against the workpiece by pressure and can move sideways along its translational degree of freedom. Unlike a curve generator, WAGNER and FAUST split the functioning elements of the machining of the surface, the positioning of the *Shape* and the generation of the *Path*. Consequently, they can be used to compensate each other. It is e.g. possible for known inaccuracies of the tool-path shaping element and for known small positioning inaccuracies of the *Shape* with respect to the workpiece, to compensate for these already in the design stage of the tool-path shaping element. A numerical analysis tool has been written, where shape errors of the generated surface due to all kinds of tool misalignments (e.g. lateral errors along the x , y , z and r axes, *Shape* inaccuracies (e.g. caused by guidance inaccuracies) or rotations of the *Shape* around the x , y , z and r axes) can be simulated. The obtained results can be used to compensate e.g. by adjusting the *Path*. Besides, for FAUST a machining process simulation has been developed to simulate the in-process generation of the surface shape as well as the shape of the cross-section of the machining band. This simulation is based on Preston's law⁴ and uses as starting point a rectangular band cross-section with a thickness δ in the x , y plane and a spherical workpiece surface. The band is rotating along the *Shape* and has only partly contact with the surface. The loaded tubular element can move along an axis z_σ that is located in the x , z plane and has an angle σ with respect to the z axis that depends on the ratio between tool and workpiece wear¹. The ratio between ω_1 and ω_2 is assumed to be constant. A point τ located on the machining band is given in the y, r, z_σ coordinate system by a vector [Eq.2]

$$\tau(n, r) = \begin{bmatrix} \cos(\sigma)(x(r) + nd \sin(\phi)) \\ y(r) - nd \cos(\phi) \\ \sin(\sigma)(x(r) + nd \sin(\phi)) \end{bmatrix} \quad \text{with: } \tan(\phi) = \frac{\frac{\partial y(r)}{\partial r}}{\frac{\partial x(r)}{\partial r}}, \quad (3)$$

$Nd = \delta$, $\Delta \leq r \leq r_m$ and $0 \leq n \leq N$ (d denotes the size of the contact area between band and surface at τ). The distances along the z_σ axis between the points τ and the corresponding surface points (the surface being machined in the same coordinate system) determine how many points along the band are worn away during one iteration step. The amount of local tool and workpiece wear is then determined, taking into account the local relative speed between band and surface together with the amount of the band surface that is in contact with the surface. That way the cross-section of the band and the surface shape being generated are simulated. This can be used to minimise the amount of *Shape* inaccuracy caused by the band thickness¹. This process simulation can also be adjusted to determine the cross-section of the wall of a small cup-wheel as it is used within WAGNER.

References

1. O.W. Föhnle, H. van Brug, C.J. van der Laan and H.J. Frankena, "Loose abrasive line-contact machining of aspherical optical surfaces of revolution," *Appl.Opt.* **36**, 4483-4489 (1997)
2. O.W. Föhnle, H. van Brug, C.J. van der Laan and H.J. Frankena, "The generation of on-axis and off-axis conic surfaces of revolution by applying a tubular tool", *Appl.Opt.* **36**, 4490-4496 (1997)
3. O.W. Föhnle, H. van Brug and H.J. Frankena, "Fabrication technique for the production of on- and off-axis conic surfaces of revolution", accepted for "Optical Manufacturing and Testing II", SPIE Annual Meeting, 1997
4. F.W. Preston, *J.Glass Technol.* **11**, 124 (1927)

Micropositioning of a Linear Motion Table with Magnetic Bearing Suspension

Oui-Serg Kim^{*}, Hyeong-Joon Ahn^{*}, Yoon Joung^{**},
In-Bae Chang^{***} and Dong-Chul Han^{****}

^{*} Graduate Student, Dept. of Mech. Design and Production Eng., Seoul National University, Korea

^{**} Production Engineering Center, Samsung Electronics Co., Suwon, Korea

^{***} Dept. of Precision Mechanical Eng., Kangwon National Univ., Chuncheon, Korea

^{****} Dept. of Mech. Design and Production Eng., Seoul National University, Korea

Abstract

This paper presents a design and performance of the 6 D.O.F. linear motion table with a magnetic bearing suspension. The linear positioning of the table with a 150 mm stroke is driven by a brushless DC linear motor and the other attitudes of the stage are controlled by the analog PD controller with magnetic bearing actuators. Each magnetic bearing unit which consists of 3 electromagnets, 3 capacitance probes and 3 backup bearings affords controlled forces by detecting the air gap between the probes and guideways.

A current integration type capacitance probe amplifier is equipped on the upper plate of the table so that the probe line to the probe amplifier can be shorter therefore the problems due to the stray capacitance and noise can be reduced.

From the pitch-yaw errors measured by the autocollimator, the vertical and horizontal straightness errors of the table are derived that they are maintained below 1 μm over 100 mm stroke. The positioning accuracy of the linear motion is maintained below $\pm 1 \mu\text{m}$ and the repeatability error is below 1 μm .

Keywords : magnetic bearing, linear motion table, straightness, micropositioning, capacitance probe

1. Introduction

In the semiconductor manufacturing systems, there are growing demands for the ultra precision equipments and especially the linear positioning system is essential. In the linear positioning system, the driving and support mechanisms determine the accuracy, so various positioning mechanisms are under way on the development. Although the magnetically levitated system with attractive type electromagnetic force is originally unstable, the feedback control plant can stabilize. The magnetic bearing is a lubricant-free system, so it can be used in the vacuum or clean room environment which can be usually encountered in the semiconductor manufacturing procedures.

In this paper, the magnetically levitated linear motion table system is developed for the probe station. The probe station analyze the wafer/substrates at the final manufacturing stage of the wafers. This station has an X-Y table, microscope, vertical positioning stage for the probes and so forth. The exchange of the already existed X-Y table into the magnetically levitated table system gives the higher positioning accuracy, reduces the friction and enables the system to be free for the contamination by lubricant. Furthermore, it can eliminate the fine attitude positioning mechanisms for the roll-pitch-yaw of the wafer, because the magnetically levitated system can

not only levitate the table but also control the fine attitude in the range of nominal clearance.

There are a few researches on this system. Trumper¹, Kuzin² and Joung^{3,4} developed the linear motion table with magnetic bearings. Tieste⁵ investigated the frequency response and Eisenhaure⁶ suggested precision design schemes on magnetically suspended table system.

This paper describes the design and performance test of a 5 D.O.F. linear motion table with magnetic bearing suspension. The 150 mm stroke brushless DC linear motor is used to drive the table and it's position is detected by the optical type linear scale. The table consists of platen and four magnetic bearing unit. Each magnetic bearing unit is composed of three electromagnets and three capacitance sensors. The performance of the table system is verified by autocollimator for the attitudes and by laser interferometer for the linear positioning.

2. System Design

Table System Fig. 1 shows that the schematics of the table system composed of six guideways, four bearing units, a table body, a linear motor, a linear scale, a sensor amplifier for twelve sensor probes and a wafer holder.

Each bearing unit consists of three electromagnets and three plate type capacitance sensor probes.

They are embedded in the bearing holder in three directions and molded as shown in Fig. 2. The specifications of the electromagnet actuator are similar to Jung⁴'s

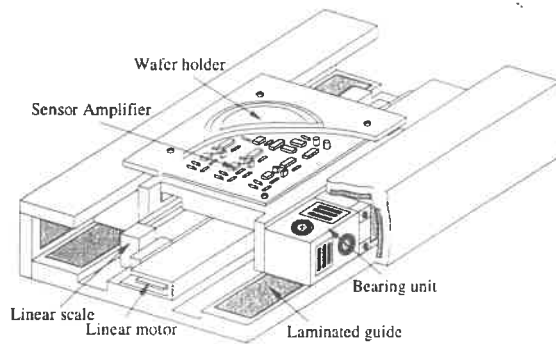


Fig. 1 Schematics of the table system

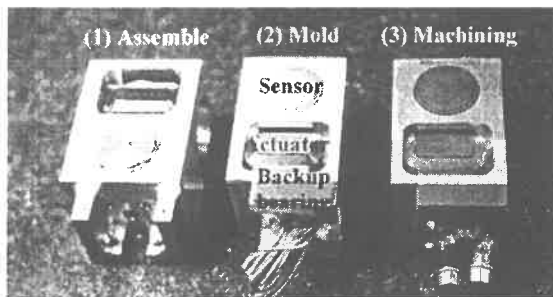


Fig. 2 Magnetic bearing units

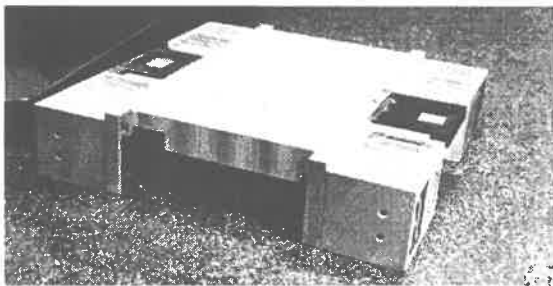


Fig. 3 Assembled table body

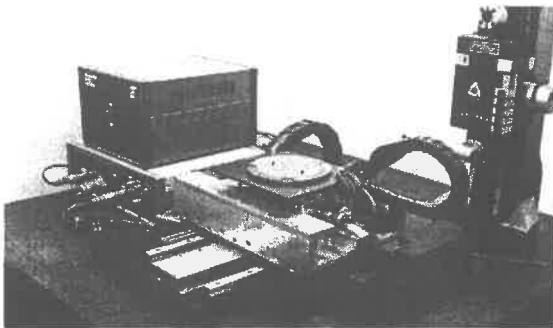


Fig. 4 X-Y micropositioning stage for probe station

Four magnetic bearing units are attached to the table body as shown in Fig. 3. The table is assembled with the guideways, linear motor and

linear scale. The guideways are made of laminated silicon steel in order to reduce the hysteresis problems of the target material. The sensor amplifier for the twelve capacitance probes is assembled on the table body to eliminate the stray capacitance and noise problems. The wafer holder is finally attached on the table as in Fig. 4. In this figure, the other linear motion axis(Y-direction) of the X-Y linear motion table system is supported by the commercial LM guide.

Sensors⁶ Twelve capacitance sensors are used to feedback the 5 D.O.F. displacement signal of the table to the controller. Configurations of sensor and guard electrodes are optimized experimentally and numerically⁷. The guard electrode enclose the sensor electrode and the triaxial cable is used so that the stray capacitance can be reduced. The two signals from the sensors at the opposite direction are electrically subtracted by operational amplifiers to get the table displacement of one direction with minimizing the baseline drift and nonlinearity. The sensor amplifier based on the charge transfer principle consists of CMOS analog switches and a current integration amplifiers. This amplifier has a 4 MHz sampling frequency and can remove the effects of stray capacitance with switched guard method. The basic circuit diagram is shown in Fig. 5 and the steady state output equation of this amplifier is very simple as given in equation (1).

$$V_{so} = -C_x V_c f R_f \quad (1)$$

Here, the C_x is the unknown sensor capacitance, V_c is the charge voltage, f is the sampling frequency, R_f is the feedback resistor.

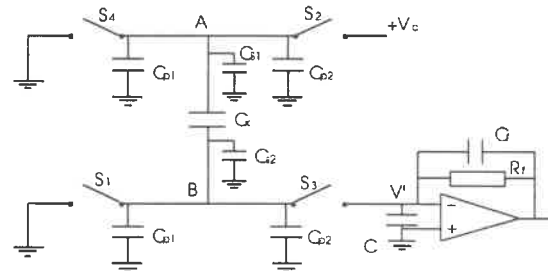


Fig. 5 Basic circuits of the sensor amplifier

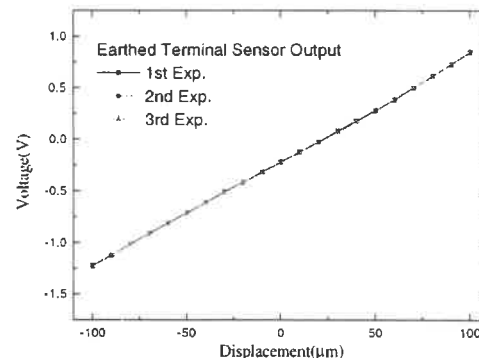


Fig. 6 The calibrated data of one sensor

Fig. 6 shows the calibrated results for one sensor probe. The linearity of this sensor/amplifier system is below 0.03 % over 200 μm range and has a 1 $\text{mV}/\mu\text{m}$ sensitivity.

Controller Fig. 7 shows the control loops of the table system. The six signals from the sensor amplifiers are transmitted to the six analog PD controllers which have offset adjustment, PD and lead compensations, low pass filter circuits and current amplifiers³. The circuit board is shown in Fig. 8.

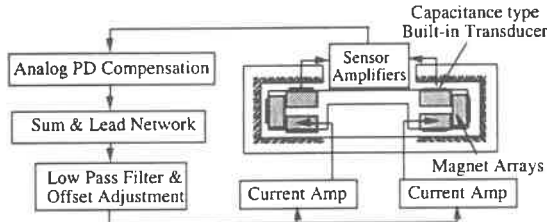


Fig. 7 Block diagram of the control system

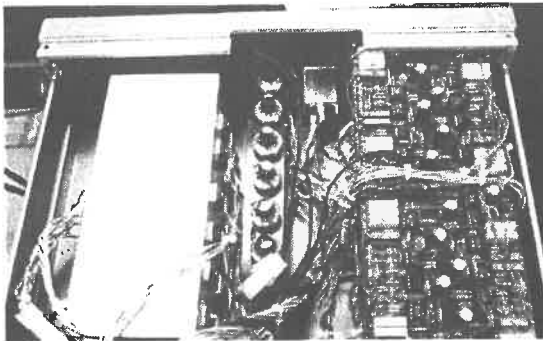


Fig. 8 Controller system with current amplifiers

Linear motor & Linear scale To get 150mm stroke of the table, linear brushless D.C. motor is used which has simple structure and has no contact nature. Normag BLC04-18 motor with 15 lbs of moving force has no attraction force in any other axis except driving axis so that the motor gives no disturbance force to the magnetic bearings. RSF MSA6703 linear scale with 10 μm of grating period shows good performance to use in position feedback system of 0.05 μm resolution.

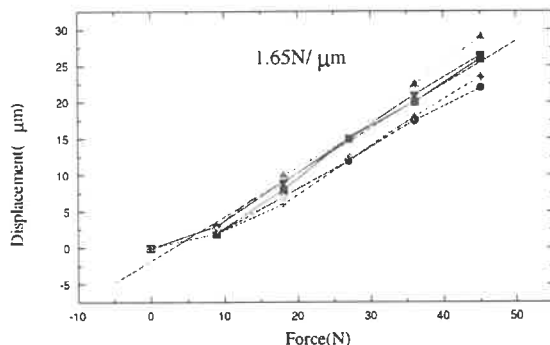


Fig. 9 The stiffness of the table

3. Experimental Results

The performances of the linear motion table are measured by three types of test schemes.

Table Stiffness The stiffness of the table is measured to 1.65 $\text{N}/\mu\text{m}$ by adding the mass at the top of the table up to 45 N range and shown in Fig. 9.

Straightness of the Table The straightness of the table is indirectly measured with the autocollimator system. The autocollimator can measure the pitch and yaw motion of the table and by using the conversion equation (2), the vertical and horizontal straightness of the table can be attained. To compensate the measured data, the incremental angular method and the least square fit is used.

$$L[\mu\text{m}] = 4.8l[m] \cdot \theta[\text{sec}] \quad (2)$$

In the equation (2), the L means the displacement of horizontal and vertical direction respectively, θ is the measured angle data and l is the travel length of each step.

Fig. 10 shows the experimental set-up and Fig. 11 shows the vertical and horizontal straightness error graphs. In the straightness error graphs, the maximum errors are remained below 1 μm over the 100 mm stroke range in each direction.

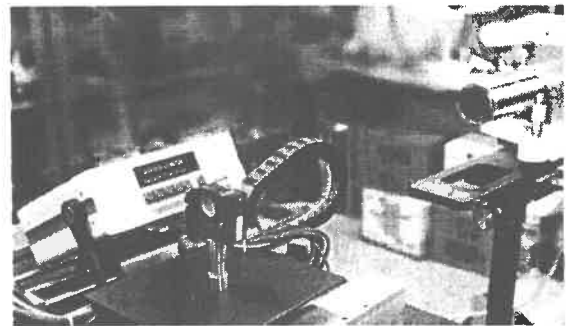
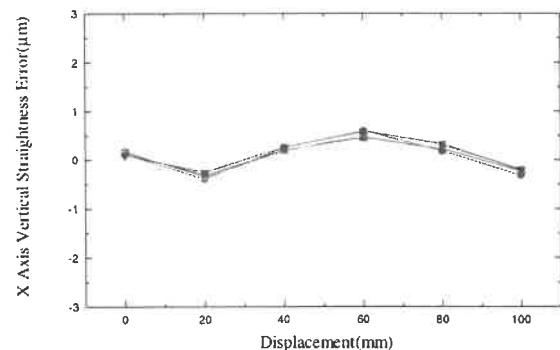
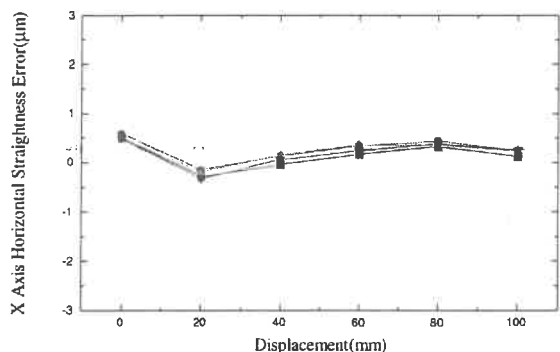


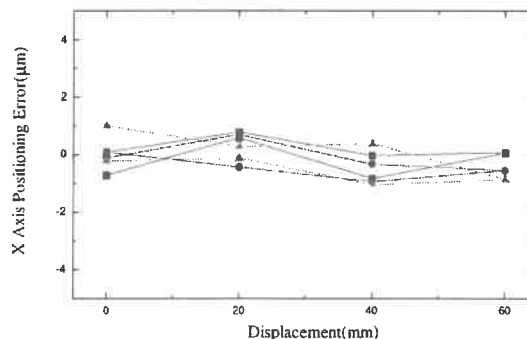
Fig. 10 Experimental set-up for straightness measurement



(a) Vertical straightness error of the table



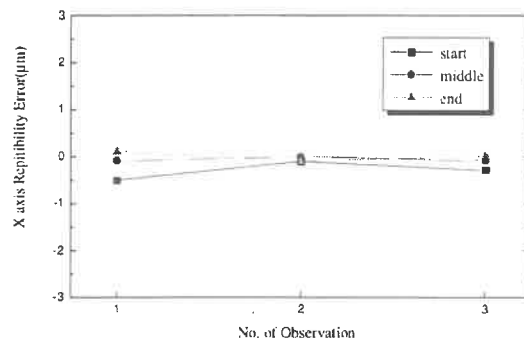
(b) Horizontal straightness error of the table



(a) Positioning accuracy of the table

Fig. 11 Straightness error of the table system

Repeatability of the Table The positioning accuracy of the table is affected not only by the linear motor system but also by the bearing support conditions and linear scale accuracy. And the accuracy of the system is measured by the laser interferometer system. The experimental set-up is shown in Fig. 12. Fig. 13 (a) shows the positioning error of the table that it is remained below $\pm 1 \mu\text{m}$ over a 60 mm range. Fig. 13 (b) shows the repeatability error of the table. The table goes forward to 20 mm and back to the initial position that the repeatability error is measured to 3 times. The repeatability error is remained below $1 \mu\text{m}$ at the both ends and middle position for the full stroke



(b) Repeatability error of the table

Fig. 13 Positioning error of the table with longitudinal direction

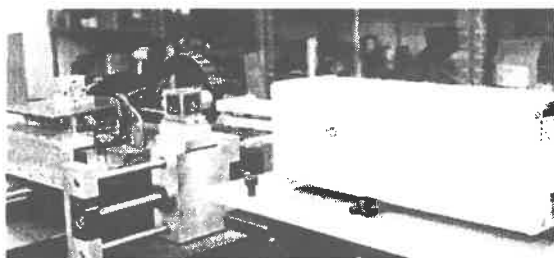


Fig. 12 Experimental set-up for the positioning error measurement

4. Conclusion

The electromagnetically levitated linear motion table system is developed for 6" wafer probe station and the performances of the table are measured.

With current integration type capacitance sensor and PD controller for magnetic bearings, advanced levitation performance can be achieved so that metrology of linear motion table is well performed.

Acknowledgments

This work was funded by Samsung Electronics Co. and supported in part by Turbo & Power Machinery Research Center and Institute of Advanced Machinery & Design in Seoul National University.

References

- [1] M.E.Williams, D.L.Trumper, "Precision magnetic bearing six degree of freedom stage", *3'rd Int'l Symp. on Magnetic Suspension Tech.*, July, 1995, pp.453.
- [2] A.Kuzin, "Precision magnetically suspended XY stage", *4'th Int. Symp. on Magnetic bearings*, 1994, 177-182
- [3] Y.Joung, S.C.Jung, I.B.Chang, D.C.Han, "An experimental study on the high-precision linear motion table with magnetic bearing suspension", *Proceedings of ASPE 10'th Annual Meeting*, Oct., 1995, pp.328.
- [4] Y.Joung, H.J.Ahn, O.S.Kim, I.B.Chang and D.C.Han "A Study on a High-Precision Linear Motion Table with Magnetic Bearing Suspension", *Proceedings of ASPE 11'th Annual Meeting*, Oct., 1996, pp.328.
- [5] K.Tieste, K.Popp, "Dynamical behavior of a linear maglev support unit for fast tooling machines", *4'th Int. Symp. on Magnetic Bearings*, 1994, 269-274
- [6] D.Eisenhaure, A.Slocum, A.R.Hockney, "Magnetic Bearings for Precision Linear Slides", *1'st Int. Symposium on Magnetic Bearings*, 1988, pp. 67.
- [7] H.J.Ahn, "A Study on a Cylindrical Capacitance Displacement Sensor for the Measurement of small rotor motion", M.D. diss., Seoul National University, 1997, Seoul, Korea

THE MACHINIST'S ROLE IN A MODERN HIGH-PRECISION MANUFACTURING ENTERPRISE

Richard L. Rhorer, Nathan R. Copeland, Charles W. Brooks, and Brian S. Dutterer
Fabrication Technology Division
National Institute of Standards and Technology, Gaithersburg, MD 20899

It has been suggested that as better systems are developed for controlling manufacturing processes, the need for the skilled machinist will decrease. The authors of this paper argue there will be an increasing role in high-precision manufacturing enterprises for people with the skills of the precision machinist.

Introduction

The role of the machinist in modern manufacturing is rapidly changing. Some manufacturing experts have suggested that in the future we will need good machine operators, but that many of the tasks traditionally handled by the skilled machinists will become automatic. An example is in the set-up of machining processes. In the past we have often relied on the experience and skill of a machinist to select the speeds and feeds best suited for a job. Tables of data were available to provide a starting point on speeds and feeds, but often a skilled machinist could greatly reduce the time and expense of optimizing a process by applying his experience to refine "book values" in particular situations. Computer Numerical Controllers (CNCs) in the future will probably have automatic speed and feed setting capabilities with appropriate sensors providing feed back from the actual process. An operator will simply have to push the "go" button and the system will adjust itself to the correct speed and feed to eliminate chatter and minimize tool wear.

As the CNCs and "virtual manufacturing systems" become better, will there eventually be no need for the skilled machinist? We think there will always be a need for the skilled machinist. And, this need for the skilled machinist will increase as industry moves to smaller lot sizes and as shops continually upgrade with new equipment. The machinist's traditional roles of making prototype parts and trouble shooting production equipment will increase, plus there will be new roles. In future manufacturing enterprises the skilled machinist's role will include participating with engineers and managers on teams which are developing new manufacturing processes. Therefore, in addition to many of the traditional skills, the next generation machinist will need some new skills.

Needed Machinist Skills

There are many things that have changed for the working machinist in the last three decades. Although machine tools have not significantly changed in appearance, even in the last hundred years, a modern precision machinist has to understand and be able to operate a machine tool through a computer based controller. Knowledge and skills of using computers, as well as the more traditional knowledge of manufacturing processes, are required by the modern machinist. We think it is very important that the "feel" that comes from making a part by hand not be lost in the training of the modern machinist. Table I lists some of the skills that we think a modern precision machinist should have as compared to a machinist of thirty years ago. In Table II, we have listed some of the skills that we think a modern precision machinist should retain even though they are not needed every day. The authors are seeking suggestions to add to the list.

TABLE I -- How Machinist's Skills have Changed

Skills taught and used 30 years ago	Skills used in modern shop
did layout work on bolt circles, etc.; used dividers, rulers, etc.; used drill press for holes	use readout on milling machine to produce bolt circle or use CNC machine
on lathes, used 4-jaw chucks, soft jaws	use 6-jaw adjustable chucks, collets
used scales and dials on milling machines to position work, had to deal with backlash	now never concerned with backlash, use readouts or CNC machines
used basic math, geometry, and trig, and had to do arithmetic quickly and accurately by hand	in addition to traditional math, use CAD/CAM (any hand calculations are done with calculator)
dimensions were always in inches	now mix metric and English units, and must know ANSI Y14.5
performed handwork using chisel, file, scraper at times to produce precision surface	only handwork done now is deburring
there was little need for communication skills	now need to read technical reports, be able to write about complicated machines and processes, be able to give good lectures and lead discussions
needed to know how to measure complicated parts with hand measuring tools (mikes, height gages, bore gages, etc.)	now need to be able to use CMMs in addition to using many of the traditional hand measuring tools
needed to make hand sketches	now use computers to make drawings
needed knowledge and skill to cut special threads on a lathe by changing lead screw gears	now use CNC lathes and program any thread, need programming skills

TABLE II -- Skills to Maintain

In addition to all the modern skills needed we recommend that the precision machinist of the future be taught the following skills that were common in the past but are in danger of being lost in the future.

- be able to center a part in a four jaw chuck
- be capable of producing a flat surface with a hand file
- be able to accurately position a slide that has backlash
- be capable of making hand drawings and sketches
- cut special threads on a manual lathe by changing lead screw gears
- modify previously machined parts, machine parts in place
- indicate vise square with machine motions, tram machine spindle

Machinist's Role in New Equipment Installation

Even more in a modern shop than in times past, new machine tools have to be installed and brought up to speed very rapidly. In a large shop a "new machine department" may handle all the details and the production department simply takes over a turn-key operation. In many smaller shops, and in research shops like our shop at National Institute of Standards and Technology (NIST), the machinist plays a significant role in the installation of a new machine tool. For example, we have installed a new high speed milling machine at NIST and one of the authors (Brian Dutterer) has been involved since the very early stages of this project. This new machine was purchased for doing machining process research and to provide a new support service capability for our shop. As the lead machinist on this project, Brian participated at every stage of the project. The details of the machinist role are described in Table III. This project proceeded smoothly and achieved its goals in a large part because of a skilled machinist's involvement. We recommend this approach for a shop of any size--the skilled machinist can make the project move more rapidly and smoothly.

TABLE III -- Machinist's Contributions to a New Machine Installation

Step in the "new machine tool" project	The role of a Skilled Machinist in this project
Justification of new machine tool	a. help define current work and jobs that are difficult or impossible: small hole drilling, thin wall, high speed b. define advantages of 4-axis machine
Preparing specifications for the new machine tool	a. check specifications for air and electrical requirements b. read and review specifications
Ordering the new machine tool	a. consider what options would make the machine more useful in our shop b. help review bids
Specify tooling	a. research tools needed for jobs intended for the machine and talk to tool suppliers b. help order tooling
Delivery of the new machine tool	a. aid in preparing utilities and new area for machine b. answer questions regarding new machine needs
Installation of the new machine tool	a. aid in proper placement of machine tool b. monitored installation for potential problems with machine c. aid in leveling and machine start up
Acceptance testing of the new machine tool	a. set-up fixtures for ball-bar test and help perform ASME B5.54 tests (by participating in these tests gain knowledge that will allow monitoring of accuracy during life of machine) b. answer questions and report results of acceptance testing to management
First parts fabricated on the new machine tool	a. prepare fixtures, blanks and part programs for some of the first parts from machine b. measure parts
Develop an efficient system of making parts in a job shop environment with the new machine tool	a. work with management to develop a scheduling approach for new machine b. explain new machine to high level management and visitors

Communication Skills and Certification

Another vital skill for the modern machinist is the ability to communicate with a broad range of fellow workers and customers. The machinist has to work in a more diverse workplace than ever before--diverse not only in working with people of different races and cultural backgrounds, but diverse in constantly changing work place situations such as new equipment and new experimental management approaches. For example, historically the precision machinist might learn from one mentor, but now the precision machinist is more likely to learn the basics in the environment of a community college, and work under many different managers during his career. To be successful he or she has to not only know how to make the parts, but must be able to tell others how the parts were made, and communicate with a variety of scientists and engineers concerning requirements and methods of fabrication. The precision machinist or instrument maker of the future will come much closer to the professional levels, which suggests that there may be a benefit for a "certification" program.

The certification of engineers, such as Professional Engineers (PE) licenses issued by states, usually involves passing a written test. For certifying precision machinists (or we would propose the title "Instrument Maker") there will be a need for some form of hands-on testing, in addition to some form of written test. An idea for a hands-on certification test could be adapted from the approach used by the Vocational Industrial Club of America (VICA) in the skills competition events. Contestants are given a drawing of a part and then allowed a given length of time to make the part. The contestants are graded on the number of features of the part that are within tolerance. In a similar vane, for a certification test, the actual fabrication of a special part could be required. This would be in addition to some form of written test verifying advanced skills, such as metrology.

We view the idea of having a certified machinist, or certified instrument maker, to be an advantage. A certified instrument maker could be relied upon to have the necessary skills of a traditional machinist plus have the advanced skills needed for working in the field of precision engineering.

Conclusions

The role of the precision machinist has changed in the past thirty years, and will continue to change. We propose that in the future the precision machinist will:

1. be a key member of teams developing and implementing new technology,
2. take a leadership role in acquiring and implementing new machine tools,
3. need to acquire computer expertise and advanced communication skills, and
4. need to work hard at maintaining traditional "hands-on" skills.

If the American Society for Precision Engineering (ASPE), or any other organization implements a certification program, we suggest that the program include a "hands-on" component as part of the certification criteria. We propose using the title "Certified Instrument Maker" if such a program is instituted.

Acknowledgments

We include a memorial recognition of Donald S. Blomquist. As the Chief of the Automated Production Technology Division, Don was a strong advocate of the project to install the new high speed mill used as an example in this paper. This machine had just recently become operational prior to Don's accidental death on July 3, 1997.

Self-Calibration of Interference Microscopes

Masaya KANAI, Wei GAO, Ichiro OGURA, Satoshi KIYONO

Department of Mechatronics, Faculty of Engineering

Tohoku University, Sendai, 980-77, Japan

ABSTRACT

The present study introduces a self-calibration method that does not require references or additional sensors to detect derivative function in order to yield a calibration curve for an interference microscope. Two specimens were used to realize the self-calibration technique. One specimen was an inclined flat surface that produces a continuous flow of calibration values. The calibration curve was derived from the difference between two measurements obtained before and after a small shift in the Z-direction. The other was a step specimen. The derivative function can be evaluated from the data obtained from the upper and lower regions of the step specimen. The results of the two specimens showed good agreement, suggesting that the proposed self-calibration technique provides an effective and reliable means of calibrating an interference microscope.

1. INTRODUCTION

Interferometric microscopes have long been used in quantitative characterization of a surface structure [1]. As a result of recent advances in phase detection techniques, interference microscopes can achieve extremely high vertical resolution [2]. However, the accuracy of interference microscopes is affected by some inherent systematic errors [3]. In particular the linearity error (deviation from a linear calibration line) of the microscope output is much larger than the resolution level. To assure accuracy of interference microscopes, the linearity error must be calibrated accurately [4]. In addition, due to the variable characteristics of linearity error with respect to the measurement conditions, the microscope should be calibrated in situ frequently to address changing conditions. The conventional calibration method requires an accurate calibration reference which is difficult to be obtained in practice. Thus, in situ calibration cannot be easily performed using the conventional method.

A self-calibration method that does not require a calibration reference has been proposed for displacement or angle sensors [5 - 7]. However this method is too large to apply to a calibration system for a restricted space because a lever system that multiplies the input to the reference side sensor is required in this method. Thus, this method is not appropriate for in situ calibration of an interference microscope.

The present paper proposes a new self-calibration method that has been specifically designed for use with an interference microscope. The proposed method does not require an extra lever system or partner sensor and can be easily applied in situ calibration. Calibration results for a commercially available interference microscope [8] are also presented.

2. PRINCIPLE

Figure 1 shows a schematic diagram of the calibration target, which is a scanning white light interferometer that employs a new spatial frequency domain analysis technique [2].

After passing through an objective lens, the light from the white light source is divided into two beams by a beam splitter in a Mirau-type objective. The transmitted beam propagates to the test surface, and the reflected beam propagates to an internal reference surface. These beams are then reflected back to the beam splitter and interfere with each other. The interference beam is received by a CCD video camera in which

each pixel corresponds to a small area of the surface.

The test surface is scanned by moving the objective in the Z-direction using a piezo-electric transducer (PZT). As the objective scans, the optical path difference (OPD) of the interferometer is varied, and the sequence of interference intensity data frames are detected and stored by the CCD camera. The data of each pixel, which represent the variation in intensity as a function of the scanning position, are transformed into the frequency domain by Fourier analysis. The relative surface height of each point of the test surface can then be obtained by calculating the phase slope with respect to the angular wave number [2].

Because the sampling position is decided by the scanning PZT, nonlinear motion of the PZT will greatly affect the linearity of the interferometer. For example, although this microscope has a vertical resolution of 0.1 nm with a maximum vertical range of 100 μm , the assured accuracy is only 1.5 % of the vertical range [8].

Assume the calibration curve of the interferometer is described by function $f(z)$ as shown in Figure 2. $f(z)$ can be generally expressed as the sum of a straight line with a slope of 1 that connects both ends of the calibration range and the deviation $g(z)$ from the straight line :

$$h = f(z) = z + g(z) . \quad (1)$$

Here, $g(z)$ is called the linearity error. Let the displacement output of the interferometer at sampling position z_i be m_i , and that at position $z_i + \Delta z$ be m_{i+1} . If the inverse function of $f(z)$ shown in Figure 3 is given by

$$h_i = m_i ,$$

$$f^{-1}(h_i) = h_i - b(h_i) , \quad (i = 1, 2, \dots, n) \quad (2)$$

where n is the sampling number, the derivative of $b(h)$ at the discrete i -th sampling point can be written as

$$b'(h_i) = 1 - \Delta z / \Delta m_i , \quad (3)$$

where $\Delta m_i = m_{i+1} - m_i$.

If Δz is given, $b(h)$ can be obtained by integration of $b'(h_i)$. If Δz is not given, Δz can be calculated as follows:

$$\Delta z = \{1/(ns)\} \sum \Delta m_i (z_{i+1} - z_i) / (z_n - z_1) \quad (4)$$

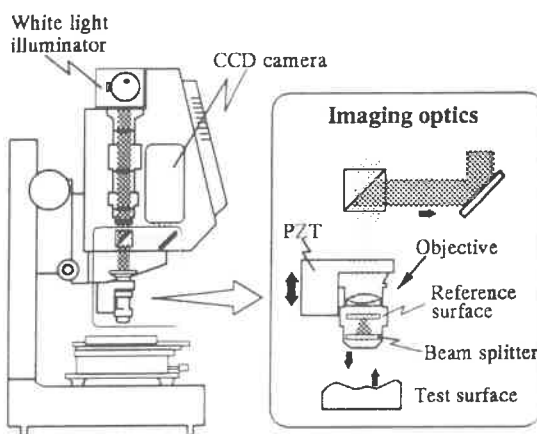


Figure 1 Schematic of microscope

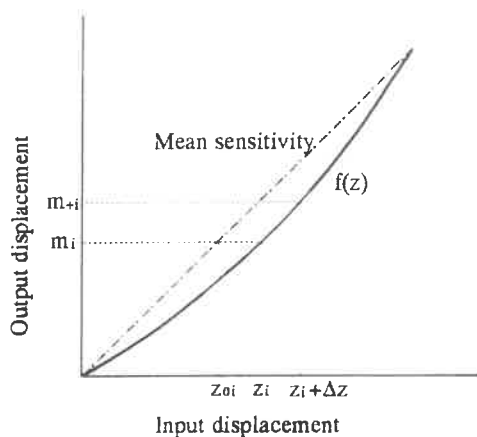


Figure 2 Calibration curve

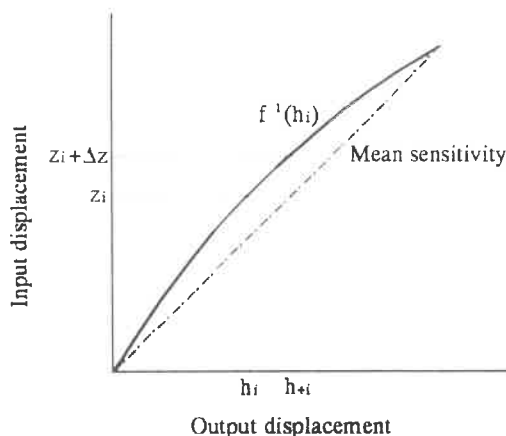


Figure 3 Inverse function of calibration curve

3. EXPERIMENT

3.1 Methods to obtain m_i and m_{+i}

Two methods are employed to obtain m_i and m_{+i} .

In Method 1, an inclined flat surface, which can give a continuous change in the input value z_i over the calibration range from z_1 to z_n , is used as the specimen as shown in Figure 4. If the inclined specimen is shifted a small distance Δz in the Z-direction, the same Δz can be applied to all points of z_i . Therefore, m_i and m_{+i} in Equation (3) can be obtained from the data measured before and after Δz .

As shown in Figure 5, Method 2 employs a step specimen, for which the step-height is assumed to be Δz . In this case, m_i and m_{+i} can be obtained from the data acquired from the upper and the lower regions of the step specimen, respectively. The evaluation accuracy can be improved by averaging the data of the upper or lower regions. To obtain m_i and m_{+i} over the calibration range from z_1 to z_n , the step specimen must be moved n times in the Z-direction with an interval of approximately Δz . Measurement is repeated at each position.

3.2 Experimental results

The calibration results obtained using Methods 1 and 2 are shown in Figures 6 (a) and (b), respectively.

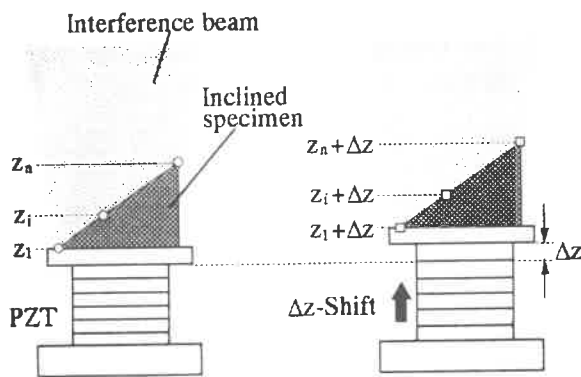


Figure 4 Principle of Method 1

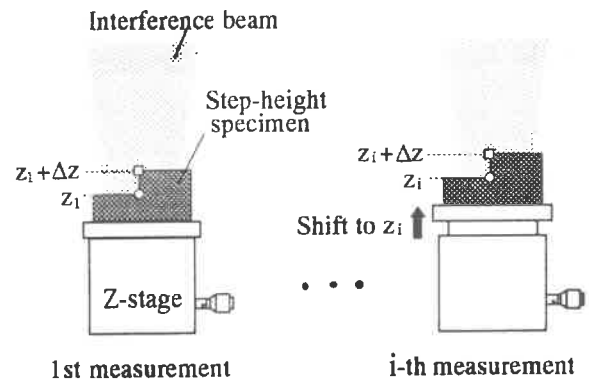
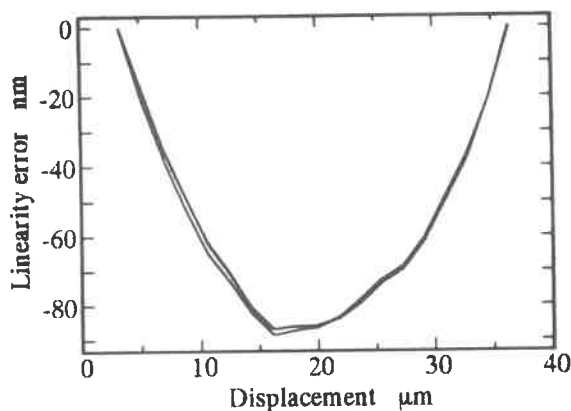
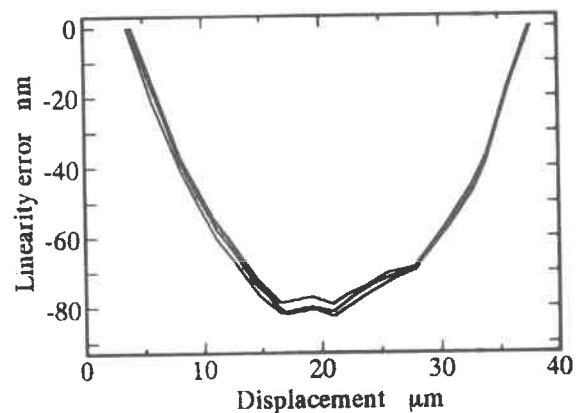


Figure 5 Principle of Method 2



(a) Results of Method 1



(b) Results of Method 2

Figure 6 Calibration results by two methods

The abscissa represents the position of the specimen in the Z-direction, which is calculated from the output of the interference microscopes. The results of three repeated calibrations are shown in each figure. The repeatability error of the three measurements in Figure 6 (a) was approximately 3 nm, and that in Figure 6 (b) was approximately 6 nm. Figure 7 shows the results obtained by fitting the average values of the three results with a fourth-order polynomial. The repeatability error, which was estimated by the root-mean-square value of the deviation associated with each measurement result from the fitting curve, was approximately 2.5 nm for each of the methods. The difference between the two fitting curves of the two methods was approximately 8 nm. These results confirmed the reliability of the proposed calibration method. The linearity error was approximately 80 nm, which is approximately 0.2 % of the calibration range of 40 μm .

In Figure 7, the derivative of each curve, for which the variation yields an error estimation of small step-height measurement, is also shown as a percentage. This figure shows that the error of the small step-height measured at different positions in the Z-direction varies from -1.2 % to 1.1 % of the step-height. This indicates that measurement data should be compensated using the calibration data in order to ensure high, accurate measurement.

4. CONCLUSIONS

To eliminate the influence of the linearity error of an interference microscope, we proposed a new in situ self-calibration technique that can accurately obtain the calibration curve by evaluating the derivative function using only the profile measurement data. Results from an inclined specimen and a step specimen have demonstrated that the proposed method is highly reliable and very effective for in situ self-calibration of an interference microscope.

ACKNOWLEDGMENT

This study was funded by a grant-in-aid for Scientific Research of Monbusho of Japan.

REFERENCES

- [1] J.H.Bruning, D.R.Herriott, J.E.Gallagher, D.P.Rosenfeld, A.D.White, and D.J.Brangaccio : Digital Wavefront Measuring Interferometer for Testing Optical Surfaces and Lenses, *APPLIED OPTICS*, Vol.13, No.11, (1974), pp.2693-2703.
- [2] P.Groot and L.Deck : Three-dimensional imaging by sub-Nyquist sampling of white-light interferograms, *OPTICS LETTERS*, Vol.18, No.17, (1993), pp.1462-1464.
- [3] J.Schwider, R.Burow, K.-E.Elssner, J.Grzanna, R.Spolaczyk, and K.Merkel : Digital wave-front measuring interferometry: some systematic error sources, *APPLIED OPTICS*, Vol.22, No.21, (1983), pp.3421-3432.
- [4] Yeou-Yen Cheng and James C.Wyant : Phase shifter calibration in phase-shifting interferometry, *APPLIED OPTICS*, Vol.24, No.18, (1985), pp.3049-3052.
- [5] S.Kiyono, K.Morisima, T.Sugibuchi : Self calibration of the linearity error of displacement sensors, *Proc. of ASPE 7th Annual Meeting, Florida*, (1992), pp.305-308.
- [6] S.Kiyono, S.Zhang : Self-Calibration of Precision Angle Sensors, *Journal of The Japan Society for Precision Engineering*, Vol.60, No.11, (1994), pp.1591-1595, (in Japanese).
- [7] S.Kiyono, Z.Ge : Subnanometric calibration of a differential interferometer, *Precision Engineering*, Vol.19, No.2/3, (1996), pp.187-197.
- [8] New View 100 Operation Manual OMP-0348K, Zygo Corporation.

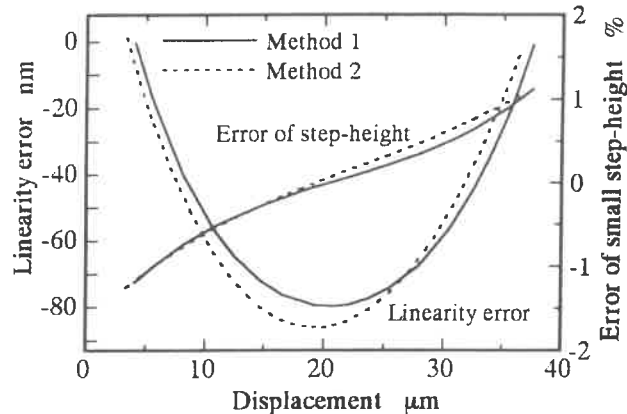


Figure 7 Averaged calibration curve

A Model for Visualization of Roller Bearing Error Motion

Wolfgang Holzhauser
The Timken Company
Canton, OH

Introduction

A simple mathematical model of tapered roller bearing error motion is described. The model simulates a bearing with an outer raceway, inner raceway and three tapered rollers equally spaced. The rolling contact surfaces of the raceways and rollers are treated as circular cross-sections with variable, sinusoidal deviations from perfect roundness. The model treats each of the elements as a rigid body, and computes the radial and axial motion of the axis of rotation of the bearing as the bearing is rotated. The goal of the model is to use very simple inputs and modeling conditions to demonstrate the deterministic nature of rolling bearing error motion, and to demonstrate that these simple inputs can nonetheless result in motions that appear extremely complex. The model has also proven of interest in understanding fundamental aspects of rolling element bearing error motion as discussion points for the next revision of the ANSI/ASME B89.3.4 Axis of Rotation Standard¹, which is currently in progress.

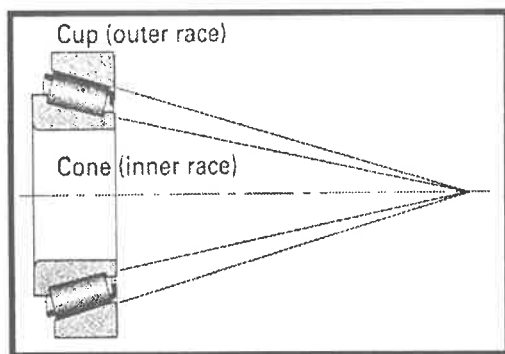


Figure 1 Basic Elements of a Tapered Roller Bearing With Common Apex Point

Tapered Roller Bearing Concepts

A tapered roller bearing consists of a conical inner and outer raceway, tapered rollers and a cage, which is used to separate and retain the rollers. As shown in Figure 1, pure rolling contact at the raceway and roller body surfaces is achieved by designing the bearing such that all of the conical elements share a common apex point. The inner ring also incorporates a conical thrust rib, which operates in contact with the spherical roller ends. A modern tapered roller bearing uses close to the maximum number of rollers that can physically fit around the periphery of the bearing, and the geometry of the internal contacting surfaces of the bearing includes features such as precision profiles and tightly controlled surfaces finish. None of these features of an actual bearing are considered in this basic model.

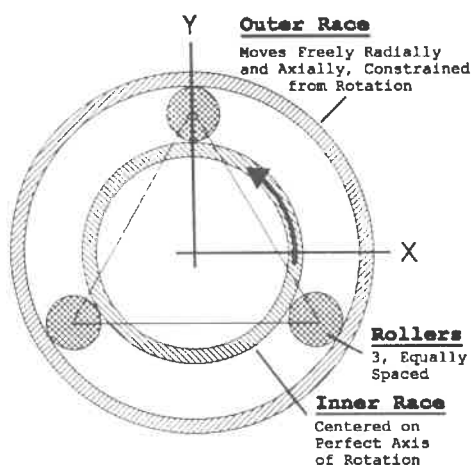


Figure 2 Model Bearing with Three Rollers

Model System

Figure 2 shows a diagram of the simplified bearing considered in the model. The inner race is rotated about its geometric centerline, and the outer race is constrained from rotation, but allowed to move radially and axially. Theoretical deviations from perfect roundness are applied to each of the rolling contact surfaces (inner race, outer race, three roller bodies). In the work presented here, the roundness deviations are single wavelength sinusoids. These roundness deviations will result in variations in 2-point diameter for the rollers, and in variations in radius for the raceways. By rotating the model bearing one angular step at a time, the resulting motion of the outer raceway as a result of these roundness deviations can be calculated with straightforward trigonometric relationships.

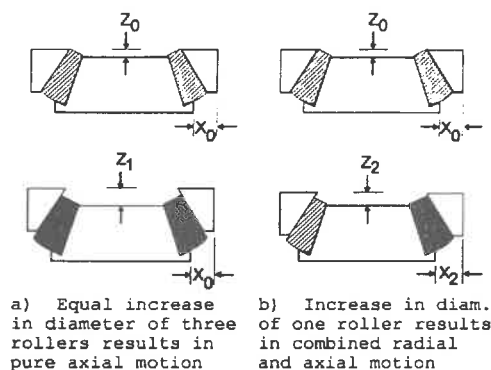


Figure 3 Axial and Radial Motions Caused by Instantaneous Changes in Diameter of (a) All or (b) One Roller. Larger Roller is Identified by Darker Shading.

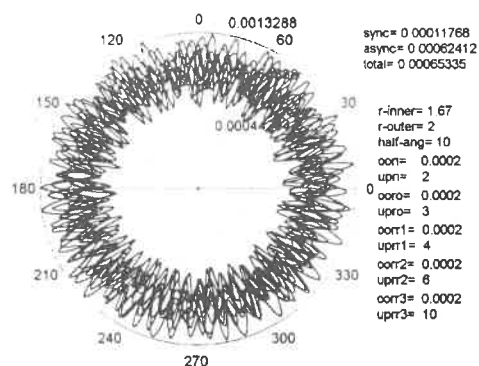


Figure 4 Error Motion Polar Plot for 10 Revolutions

evident that the non-rotating element of the bearing undergoes a variety of motions, including large excursions and rapid reversals of direction.

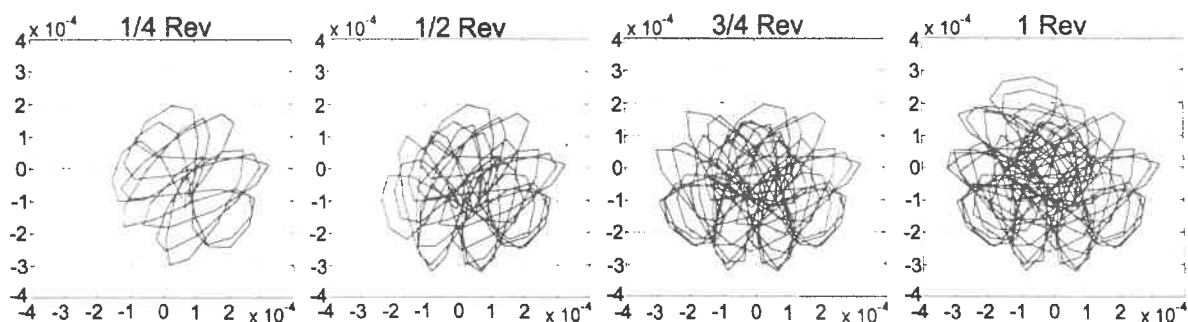


Figure 5 Error Motion XY Plot for 1/4, 1/2, 3/4 and 1 Revolution

While the radial motions represented in Figures 4 and 5 are taking place, axial motion also occurs in this model bearing. Figure 6 shows an XYZ plot of the trajectory of a fixed point on the outer race for one revolution of the inner race. The Z axis displacement is the result of sliding motion at the roller body/raceway contact

A unique characteristic of a tapered roller bearing is the coupling of radial and axial forces and/or motions as a result of the tapered design. For example, if the instantaneous diameter of all three rollers in the bearing increases, axial motion of the outer race relative to the inner race can occur to accommodate the resulting geometric interference. This is shown in Figure 3a where dimension $Z_1 > Z_0$. In Figure 3b, an increase in instantaneous diameter of one roller will result in both radial and axial motion ($X_2 > X_0$ and $Z_2 > Z_0$).

Results

Figure 4 shows the error motion polar plot for a single, fixed sensitive direction for an example case in which the inner raceway mid-plane diameter is 1.67 units, the outer race mid-plane diameter is 2.00 units, the inner race has a 2 undulation per revolution (UPR) roundness deviation, the outer race has 3 UPR roundness deviation, and the rollers have 4, 6 and 10 UPR deviations respectively. All of the roundness deviations have a magnitude of 0.0002 units. Note that the units used are arbitrary: they can be inches, millimeters, etc. The plot represents 10 revolutions of the inner race. The synchronous, asynchronous and total error motion values are also shown on the plot.

The single sensitive direction polar plot in Figure 4 is the standardized [1] method of representing error motion data, and gives an excellent indication of the potential performance of a bearing or spindle on some machine tools or measuring instruments. However, it does not represent the complete motion of the bearing outer race relative to the rotating inner race. For this, the XY plot in Figure 5 is more useful. This plot shows the trajectory of a fixed point on the non-rotating outer race for 1/4, 1/2, 3/4 and 1 revolution of the inner race. It is

(assumed 10 degree half angle relative to bearing axis). Note that in this simple model with no inertia or sliding friction, the magnitude of axial motions can be quite large relative to the radial motions.

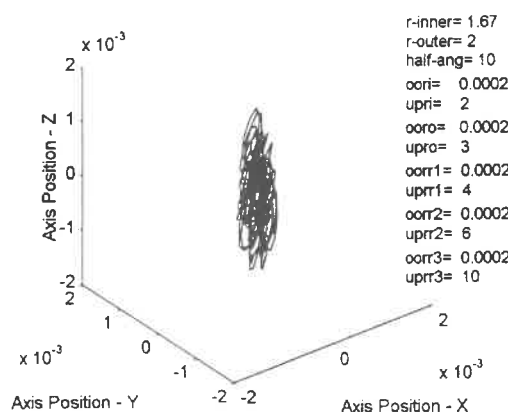


Figure 6 Error Motion XYZ Plot for 1 Revolution

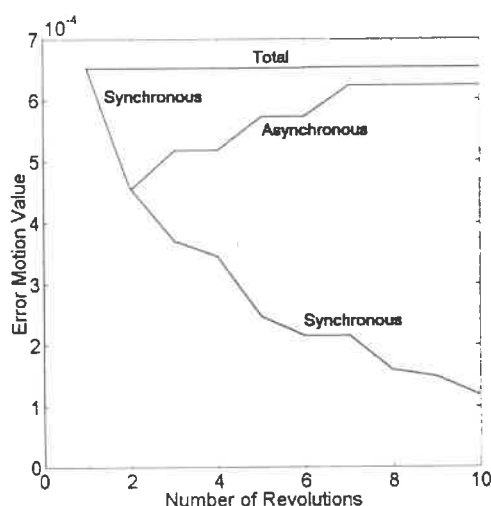


Figure 7 Error Motion Values Versus No. Of Revolutions

The model also shows behavior that has been observed experimentally on rolling element bearings. One example is the development of values of synchronous, asynchronous and total error motion values shown in Figure 7. Note that after the first two revolutions, the asynchronous error motion value increases, but that this increase sometimes occurs in steps. This phenomenon is also observed experimentally and can be explained by considering the definition of the asynchronous error motion value. This value is dependent on the largest excursions seen on the error motion polar plot. If during one or more revolutions of the bearing no excursions occur that are larger than those observed already, the asynchronous error motion value will remain constant. A stepwise increase then occurs in the subsequent revolution in which a larger excursion occurs. Also note that no asynchronous error motion value is plotted for 1 revolution, as the concept of asynchronous motion makes no sense for a single revolution (the same could be argued for the synchronous and total error motion values).

Figure 7 also reveals an interesting characteristic of a tapered roller bearing with three rollers equally spaced. As the number of revolutions increases, the synchronous error motion value decreases. Additional work with the model and with actual bearings has shown that the synchronous error motion value approaches zero after a sufficient number of revolutions. This occurs, however, at the expense of generating high values of asynchronous error motion.

Real Time Visualization

One of the most useful aspects of the model is in viewing the development of the error motion plots as the model bearing goes through it's simulated motions. A particularly useful representation of the data is to build the error motion XY or XYZ plot continuously, and display the most recent few data points in a different color from the rest of the plot. This method shows a short-term pattern driven by certain aspects of the part geometric deviations, and this pattern in turn

moves around the plot area. The relative frequencies of these motions are determined by the bearing characteristic frequencies, and by the wavelength selected for the roundness deviation of each surface.

Fundamental Motions

The relatively complex-appearing error motion plots in Figure 4, 5 and 6 are actually comprised of some very basic motions. These motions are:

- Circular precession in the XY plane due to the 2 UPR component on the inner race.
- Axial motion in the Z direction, with zero motion in the XY plane, due to the 3 UPR component on the outer race.
- "Daisy" shaped patterns of three different frequencies in the XY plane, plus synchronized axial motions, due to the 4, 6 and 10 UPR roundness deviations of the three rollers.

These basic motions in XY and XYZ are plotted in Figures 8 and 9.

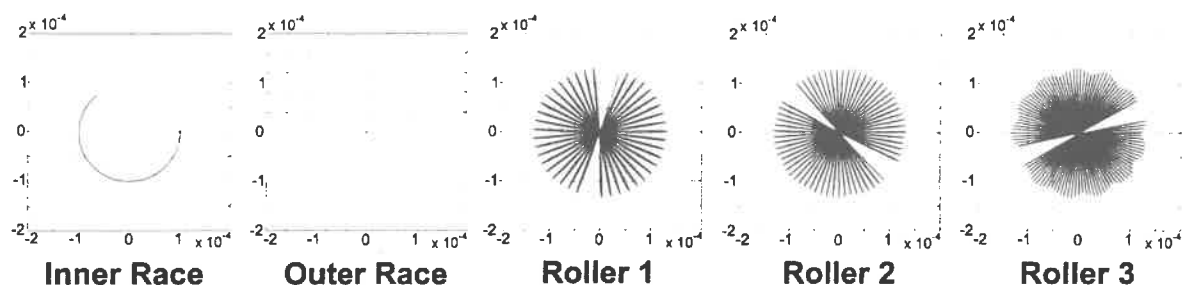


Figure 8 Error Motion XY Plots Showing Motions Generated by Each Bearing Element

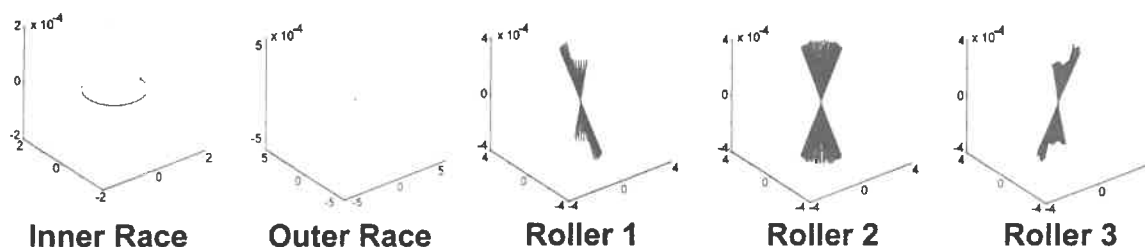


Figure 9 Error Motion XYZ Plots Showing Motions Generated by Each Bearing Element

Note in Figure 8 that the furthest right plot contains a false lower frequency which results from the selected angular increment in the calculations (1 degree angular increment) and the relatively high frequency of the motion. It is to be expected that the model reveals the same dangers of digitally sampling data as are encountered when sampling real data. The angular increment used in the model can be adjusted to minimize these effects if desired.

Conclusions

This model is clearly highly idealized and also restricted in the types of error motions that can be generated. The model does not consider tilt motion, the effect of the inner race thrust rib, slip, friction, inertia, component elasticity or more than three rollers. However, it has formed an excellent first step in the development of more sophisticated models that consider these factors. The simplified three roller model is also useful as a test bed for the development of new error motion parameters or concepts. Even though the idealized three-roller bearing is very simple, any generalized definitions and concepts proposed for understanding error motion should still be applicable to it. It has also proven useful for visualization of bearing error motion, and understanding the fundamental relationships between traditional single sensitive direction error motion polar plots and the XY and XYZ error motion trajectory plots shown here.

References

1. ANSI/ASME Standard B89.3.4M - 1985, "Axes of Rotation - Methods for Specifying and Testing," American Society of Mechanical Engineers, New York, NY, 1985.

GEOMETRIC ERROR CALIBRATION OF MULTI-AXIS MACHINES USING AN AUTO-ALIGNMENT LASER INTERFEROMETER

Jenq Shyong Chen

Department of Mechanical Engineering, National Chung Cheng University

160, San-Hsing, Ming-Hsiung, Chia-Yi, 621, Taiwan, R.O.C.

Tel:886-5-2428113, Fax:886-5-272-1522, e-mail: imejsc @ccunix.ccu.edu.tw

Introduction

Software error compensation technique has been proven to be a cost-effective method to enhance the machine accuracy. Volumetric error compensation of a multi-axis machine, however, requires the detailed information of machine geometric errors which are usually calibrated using laser interferometers. Commercial laser interferometer systems such as the HP5529A are expensive because it requires different optic sets for each kind of geometric error. Geometric calibration using commercial laser interferometer is also time-consuming because each error calibration requires a tedious work for initial optic setup and manual laser-path alignment. These commercial laser interferometers, designed major for single axis movement calibration, are also very difficult for a diagonal measurement where multiple axes are moved simultaneously. However, the diagonal measurement is found to be a better presentation of volumetric accuracy for a multi-axis machine.

This work will report the development and application of using an auto-alignment laser interferometer system to calibrate the geometric errors of multi-axis machines. The auto-alignment laser interferometer system is able to do any diagonal displacement measurement without the necessity of laser beam alignment. One application of the auto-alignment laser interferometer technique is that geometric errors (including the angular roll errors) can be quickly calibrated or estimated from the displacement measurements of some determined diagonal lines in the machine working zone. Some experimental results will be discussed.

System Configuration

The auto-alignment laser interferometer system is based on a Hewlett-Packard HP5529A laser interferometer system. Fig. 1 shows the configuration of this system implemented in a three-axis CNC machine. The system includes a laser head, an interferometer mirror, a retroreflector mirror, a fiber optic bundle, a lateral effect photodetector and computer interfaces. For a conventional laser interferometer system, the laser head, interferometer and retroreflector must be carefully aligned to parallel the motion line of a machine axis before calibration. In this modified system, the laser path between the laser head and interferometer is connected using a fiber optic bundle developed by POINT SOURCE LIMITED. The laser from the laser source to the interferometer is connected by a single mode and polarization preserving fiber optic. The laser from the interferometer to the receiver is connected by a step-index multimode optic fiber. Using the optic fiber, the alignment between the laser head and interferometer can be eliminated.

Laser path alignment between the interferometer and reflector is achieved using a lateral effect photodetector and computer interface techniques. A lateral effect photodetector is attached on the interferometer as shown in Fig. 2a. Laser beam reflected back from the retroreflector will return to the interferometer. Part of the returned laser will be reflected by the interferometer and be projected on the surface of the lateral effect photodetector. During measurement, the interferometer will be programmed to move along one machine axis which is called the active motion. If active motion causes any relative lateral displacement between the interferometer and retroreflector, the laser spot position on the photodetector will be changed too, as shown in Fig. 2b. The PC continuously monitors the changing of the laser spot position and informs the CNC controller to coordinate the relative lateral displacements between the interferometer and retroreflector in order to make the laser spot on the surface of photodetector back to the initial position. The coordinated relative lateral displacements between the interferometer and retroreflector are called passive motion here. Through the control of passive motion, the alignment between the interferometer and retroreflector during measurement can be maintained.

System Evaluations

The reflected light from the beam splitter of the interferometer is used to align the laser path. Because the beam splitter is coated with anti-reflection coating, the intensity of the reflected laser projected on the photodetector is very low. Therefore, the noise level or the stability of the laser spot position is concerned. It is found that the position stability of the laser position caused by noise level is within 0.01mm which is good enough for most applications. Conventional laser interferometer system using manual alignment is very difficult to achieve this level. Fig. 3 shows the measurement results along a single axis measurement. The displacement errors are measured three times in a very short time period in order to eliminate the environmental effect. It shows that the modified system gives the same or slightly better performance in repeatability than the original HP5529A. It was probably because the air-turbulence-caused deterioration of the laser properties between the laser source and interferometer mirror was eliminated by the optic fiber transmission. Demonstration of 3D diagonal measurements is presented in the Fig. 4. One application of the diagonal displacement measurement capability is the volumetric accuracy evaluation of a multi-axis machine. Because the fiber optic cable will be swung during measuring, it is concerned that the measurement stability due to the swing effect of the optic fiber cable. Fig 5 shows the measurement result of Y-axis displacement error using different feedrates (300mm/min, 1000mm/min, 3000mm/min, and 6000mm/min). It is shown that the swing effect of the cable is very limited, if the feedrate is below 3000mm/min.

Geometric Error Calibration

Another application of the auto-alignment laser interferometer system is that geometric errors of a multi-axis machine can be quickly calibrated from the diagonal measurement capability. Using this method, optic sets for angular, straightness and squareness measurements of a HP5529A system can be eliminated. Sophisticated training of optic setup and alignment is also not required, because only simple displacement measurement is needed. Geometric errors (including the angular roll errors) can be calibrated or estimated from the displacement measurements of some determined diagonal lines in the machine working zone. For example, pitch and yaw error can be calibrated from displacement measurements of parallel lines with a known offset distance between them. Fig. 6. shows the measurement results of pitch angular errors using the modified system and HP5529A system.

Other geometric errors such as straightness, squareness and roll error can also be calibrated from 2D diagonal displacement measurements at some determined planes. Squareness can be calibrated from the displacement measurements of two diagonal lines of a plane. As shown in Fig. 7, if there is no squareness error, the displacement errors of the diagonal line AB would be same as the errors of the diagonal line CD. However, the displacement errors of the line AB and CD would be different if there is squareness existed. Similarly, roll errors can be calibrated from the displacement measurements of diagonal lines located in two parallel planes. As shown in Fig. 8, if there is no roll error, the displacement errors of the line AB will be same as the errors of the line CD. However, displacement errors of the line AB and CD will be different if there are roll angular errors existed in the machine axis. Fig. 9 shows the comparison of the roll errors measured by the developed system and by an electronic level. It shows that the results are closed.

As demonstrated in the reference [1], straightness error can also be calibrated from the displacement errors of three diagonal line of a plane. However, this method involves an iteration calculation of the displacement errors. If the machine repeatability is not good, the iteration procedure could cause accumulated estimation error which makes this calibration method unreliable.

Conclusions

An auto-alignment laser interferometer system for machine tool calibration has been demonstrated. This system can do diagonal displacement measurements without the necessary of time-consuming laser path alignment procedure. Geometric errors of a multi-axis machine can be quickly calibrated from the diagonal measurement. It is expected that the geometric errors of a machine tool can be calibrated within one hour using the auto-alignment laser interferometer technique, compared with a time of several days using the traditional laser interferometer system. However, estimation of the straightness errors from the displacement measurement may be unreliable if machine repeatability is poor. In summary, benefits expected from this new technique are: 1) ease of setup and no alignment required for any diagonal measurement, and 2) single set of optics for all 21 geometric error calibration.

Acknowledgment

This work was supported by the National Science Council grant No. NSC86-2212-E-194-006.

Reference

[1]. Zhang, G., Ouyang, R., Lu, B., Hocken, R., Veale, R., and Donmez, A., "A displacement method for machine geometry calibration", Annals of the CIRP, Vol. 37/1/1988

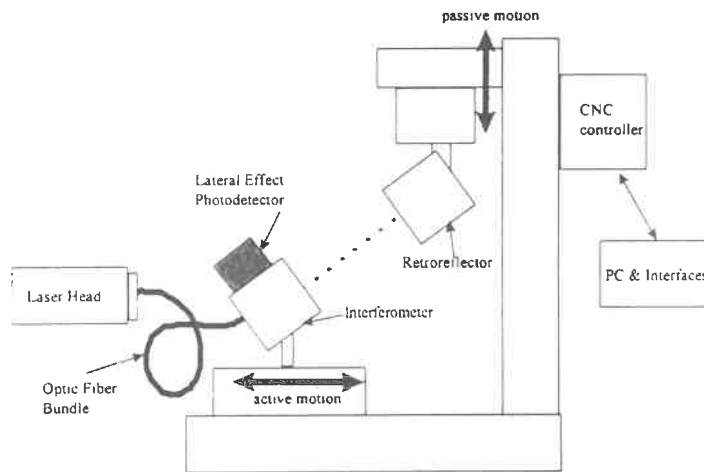


Fig. 1: The configuration of a auto-alignment laser interferometer

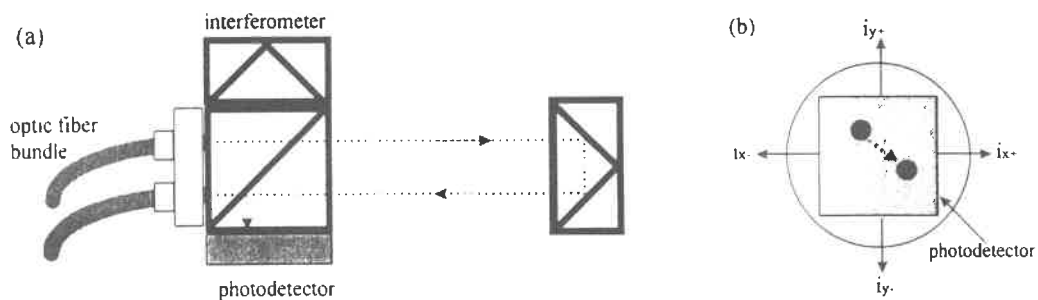


Fig. 2: Projection of the reflected back laser spot on the surface of the lateral effect photodetector

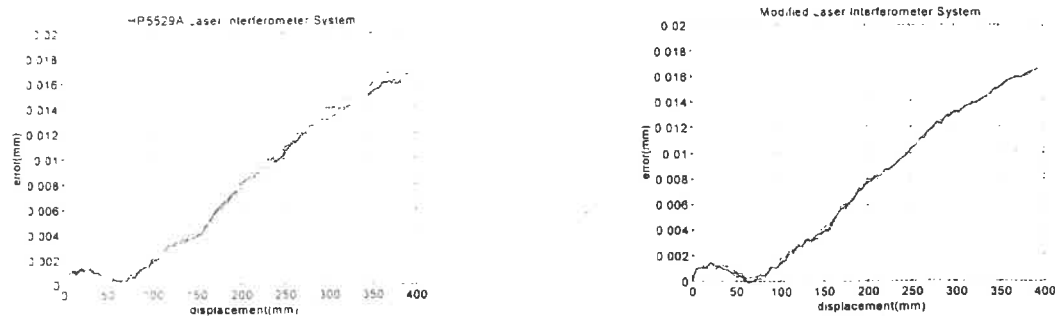


Fig. 3: Comparison of the modified system and the HP5529A

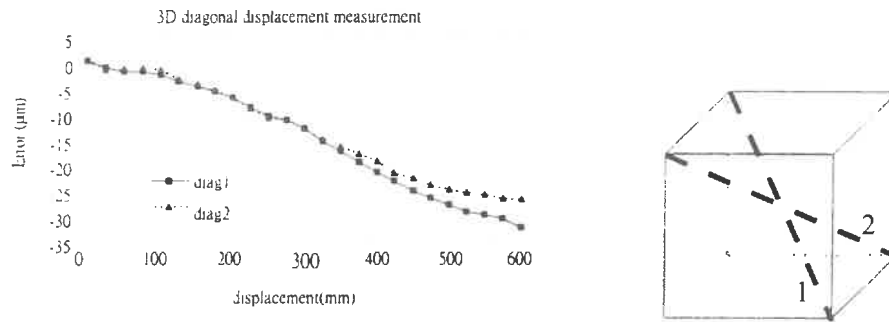


Fig. 4: Example of a diagonal displacement measurement

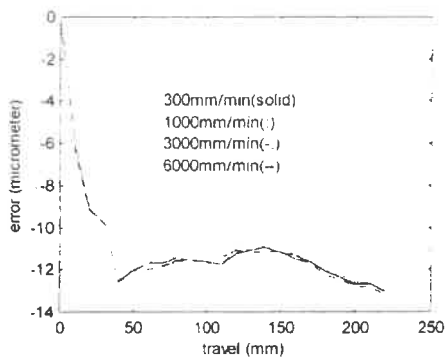


Fig. 5: Effects of machine feedrates

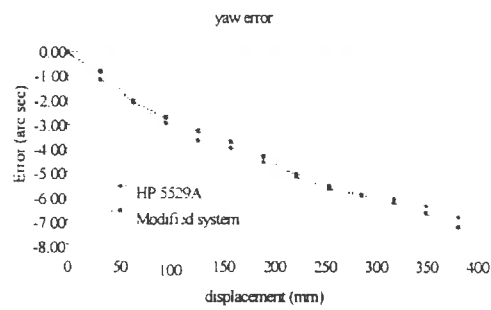


Fig. 6: Calibrated pitch angular error

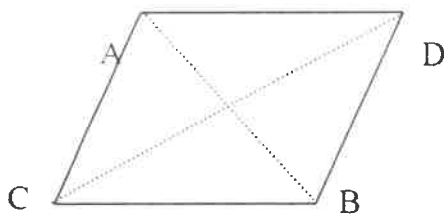


Fig. 7: Calibration of squareness

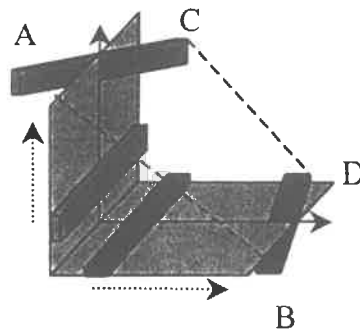


Fig. 8: Calibration of roll errors

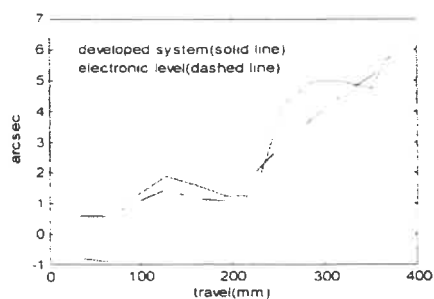


Fig. 9: Calibrated roll errors

PRECISION PROFILE MEASUREMENT OF ASPHERIC SURFACES BY IMPROVED RONCHI TEST

Seung-Woo Kim, Ho-Jae Lee
Korea Advanced Institute of Science and Technology, Department of Mechanical Engineering,
Yusong-gu, Taejeon, 305-701, Korea

Seong-Eun Lim,
Samsung Electronics, Core Technology Research Center, Suwon, 442-742, Korea

Abstract

We present an improved version of Ronchi test that is particularly intended for the precision profile measurement of aspheric surfaces of mold cores. A sinusoidal grating is used, instead of conventional rule grating, to obtain high quality lateral shear interferograms with only +1, 0, and -1 order diffraction terms. Undesirable higher order diffraction terms are inherently removed for better signal-to-noise ratio of measured interferograms. Further, the relative phases of interferograms are accurately computed by adopting phase-shifting technique along two orthogonal shear directions. Finally, the 3-dimensional profile of the aspheric surface under inspection is reconstructed using the wavefront extracted by a dual-path wavefront integration algorithm. As compared with the conventional Ronchi test, this improved version of Ronchi test provides a great improvement in measurement stability.

Keywords: Optical testing, aspheric surfaces, Ronchi test, lateral-shearing interferometer, sinusoidal gratings.

Introduction

Aspheric lenses are increasingly used in consumer products such as video cameras and compact disc pickups, as their mass production has become possible owing to recent remarkable innovations made in the ultraprecision plastic mold injection technology. However, still much work needs to be done for more effective manufacture of complex-shaped aspheric surfaces, especially in the area of precision profile measurement of fabricated mold cores and lenses. Traditionally various optical null tests have been widely used for optical surface quality assurance, but they are found not productive for aspheric surfaces due to the difficulty in preparing accurate aspheric reference surfaces. Instead, the principle of lateral-shearing interferometry draws much attention since it can operate without reference surfaces. The well-known Ronchi test is a special form of lateral-shearing interferometry in which necessary shearing is achieved by diffraction using a rule grating.[1]

Optical Setup and Measurement Theory

Figure 1 shows the optical setup of Ronchi test configured in this investigation. A collimated HeNe laser beam is focused by a microscope objective to form a point source at S_o . Then the diverging beam from S_o is directed to the aspheric test surface being measured and then reflected back through a beam splitter to converge to the point S_r that becomes the conjugate point of S_o . A sinusoidal rule grating is located at S_r to diffract the converging beam, and finally the -1, 0, and 1 order diffracted terms are imaged on a CCD array being overlapped to form Ronchigram.

The wavefront reflected from the test surface may be expressed as $w(x,y) = w_s(x,y) + w_a(x,y)$, i.e., the spherical term $w_s(x,y)$ and the non-spherical term $w_a(x,y)$. The spherical term may be explicitly written as $w_s(x,y) = (x^2 + y^2) / 2d_2$ by adopting the Fresnel approximation, where d_2 is the distance from the surface to S_r . On the other hand, the non-spherical term $w_a(x,y)$ is to be directly measured to construct the whole wavefront $w(x,y)$. At the point S_r which corresponds the Gaussian image point of S_o , the spherical term cancels out and the complex amplitude is determined as [2]

$$g(u,v) = \exp[j \frac{\pi}{\lambda d_2} (u^2 + v^2)] \int f(x,y) \cdot \exp[j \frac{k}{d_2} (xu + yv)] dx dy \quad (1)$$

where $f(x,y) = \exp[jkw_a(x,y)]$ and $k = 2\pi / \lambda$.

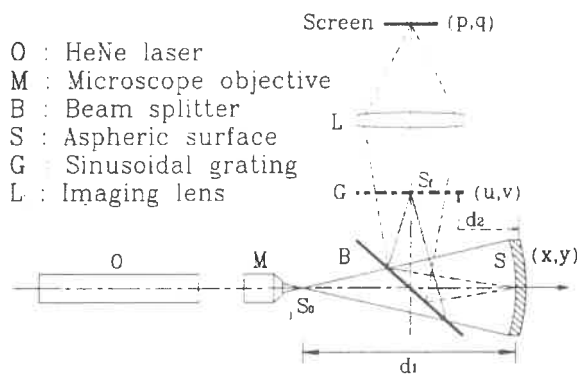


Figure 1. Optical setup for measurement of aspheric surfaces

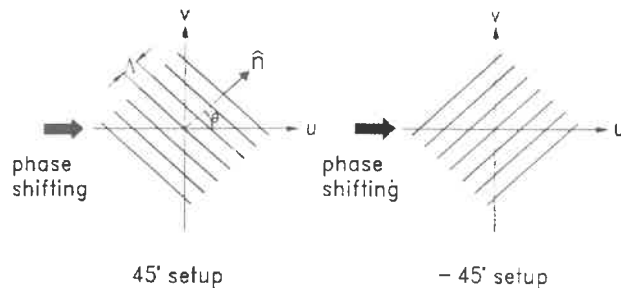


Figure 2. Alignment of sinusoidal grating

Figure 2 shows the orientation of the sinusoidal grating aligned in the uv -plane, in which Λ and $\bar{n}$ represent the pitch and the normal vector of the line grating, respectively. The transmittance of the grating is given by $t(u,v) = 1 + \cos[(2\pi / \Lambda) \bar{n} \cdot \bar{r}]$, which is obtained as

$$t(u,v) = 1 + \frac{1}{2} \left[\exp[j \frac{2\pi}{\Lambda} (u \cos \theta + v \sin \theta)] + \exp[-j \frac{2\pi}{\Lambda} (u \cos \theta + v \sin \theta)] \right] \quad (2)$$

The complex amplitude of the beam right after the grating is modulated such as $g(u,v) \cdot t(u,v)$. The CCD array is placed on the conjugate plane of the test surface, thus the intensity distribution on the CCD array becomes the inverse Fourier transform of $g(u,v) \cdot t(u,v)$, which is worked out as

$$I(p,q) = I_0 [I_{(0,1)}(p,q) + I_{(0,-1)}(p,q) + I_{(1,-1)}(p,q)] \quad (3)$$

In the above, I_0 is the mean intensity, $I_{(0,1)}(p,q) = \cos[k(w_a(p,q) - w_a(p - s_p, q - s_q))]$, $I_{(0,-1)}(p,q) = \cos[k(w_a(p,q) - w_a(p + s_p, q + s_q))]$, $I_{(1,-1)}(p,q) = \cos[k(w_a(p - s_p, q - s_q) - w_a(p + s_p, q + s_q))]$, $s_p = \lambda m d_2 \cos \theta / \Lambda$, $s_q = \lambda m d_2 \sin \theta / \Lambda$ where m is the magnification of the imaging lens. The intensity distribution in fact results from the interference occurring among the -1, 0, and 1 order

diffracted beams of Eq.(1).

Phase-shifting technique may be usefully adopted to analyze the Ronchigram of Eq.(3). If an equal amount of phase shift $2\pi/N$ is induced N times successively, the relative phase to be determined is related to the non-spherical wavefront such as [3]

$$\tan \phi(p,q) \equiv \frac{\sum_{n=1}^N I_n(p,q) \sin \Delta_n}{\sum_{n=1}^N I_n(p,q) \cos \Delta_n} = \tan \left[\frac{\partial w_a}{\partial p} s_p + \frac{\partial w_a}{\partial q} s_q \right] \quad (4)$$

where $\Delta_n = (2\pi/N)(n-1)$. Therefore, once $\phi(p,q)$ is determined from Eq.(4), $w_a(p,q)$ is then computed through an appropriate integration process. For accurate computation of $w_a(p,q)$, it is necessary to incline the grating as shown in Figure 2 and also to implement phase-shifting technique twice with $\theta = +45^\circ$ and -45° , respectively.

As the final step of the measurement, the surface profile is reconstructed from the measured wavefront $w(x,y) = w_s(x,y) + w_a(x,y)$. If it is assumed that the incident beam to the surface is in parallel with the optical axis as illustrated in Figure 3, the surface slope is related to the wavefront such as

$$\frac{\partial S}{\partial x} = \frac{1 - \tan^2 \theta_x}{2} \frac{\partial w}{\partial x} \quad \text{where} \quad \tan 2\theta_x = \frac{\partial w}{\partial x}. \quad (5)$$

As for the general case when the incident beam is not in parallel with z -axis, the relationship between S and w becomes complicated to require numerical computation. Further description is omitted here for space limitation.

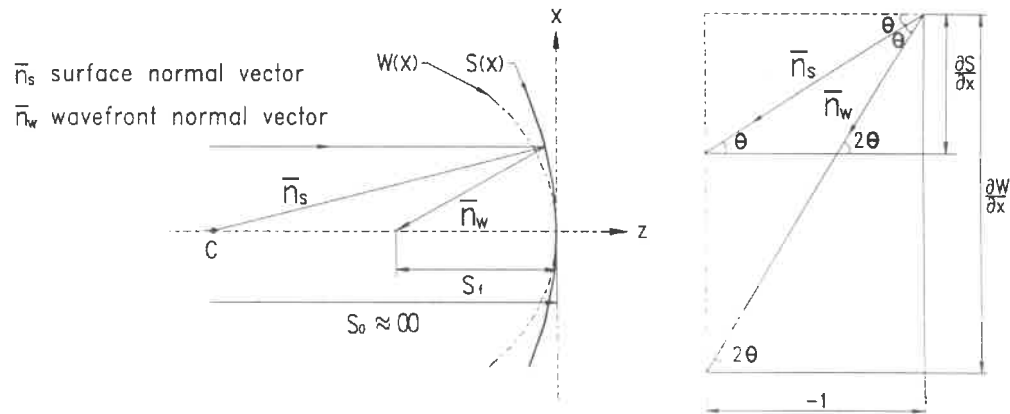


Figure 3. Relation between aspheric surface(S) and wavefront(W)

Experiment and Discussion

Figure 4 shows a typical Ronchigram and its corresponding profile of the test surface. Two different types of line gratings, one is the conventional rule grating of square transmittance fabricated by electron-beam lithography, while the other a sinusoidal grating made by holographic process on polyester film. Theory previously described predicts that the two films provide identical results, as higher order terms of diffraction than $-1, 0, 1$ do not affect the measurement. In practice, however, the square grating tends to generate more noises

deteriorating measuring accuracy. Figure 5 demonstrates a comparison of measurement results in that the sinusoidal grating provides a higher repeatability than the square grating in wavefront reconstruction.

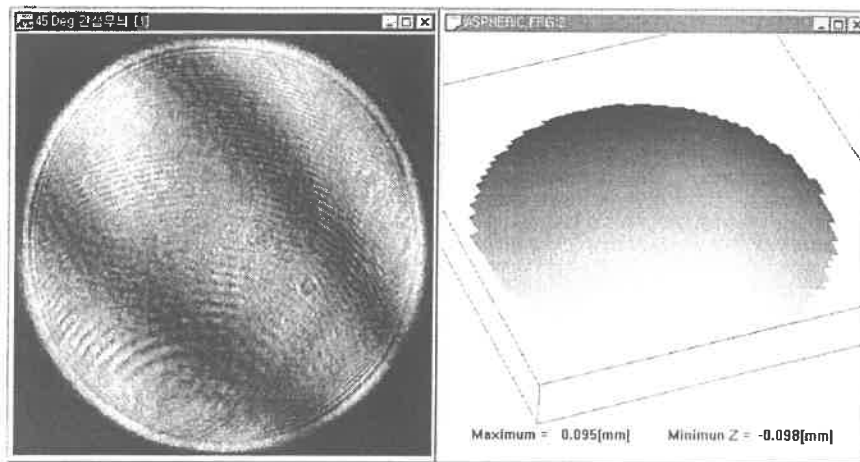


Figure 4. Ronchigram and reconstructed surface profile

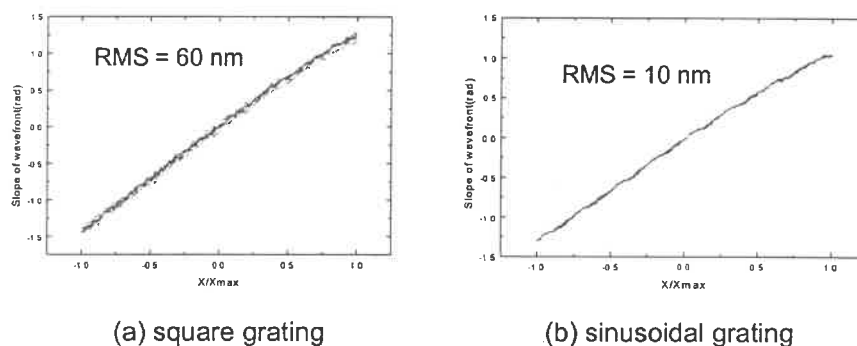


Figure 5. Reconstructed wavefront derivatives

Conclusion

An improved method of Ronchi test has been proposed with the intention of precision profile measurement of aspheric surfaces. Particular emphasis has been placed on generalization of measurement theory so that a flexibility of optical setup is gained in handling short-focal-length surfaces with large numerical aperture by using a point source laser beam instead of plane wave for illumination. In addition, a great enhancement in measuring accuracy has been achieved by compensating for the alignment errors of optical components through a well-established relationship between the measured fringe intensities and the surface profile.

References

- [1] D. Malacara, "Optical shop testing," John Wiley & Sons, Inc., 2nd edition, 1992.
- [2] A. VanderLugt, "Optical image processing," John Wiley & Sons, Inc., 1992.
- [3] K. Omura, T. Yatagai, "Phase measuring Ronchi testing," Appl. Opt., 27, 523(1988).

Surface and Form On-Machine Metrology / Correction System for Ultraprecision Machining and Grinding

Gene Malkin
Precitech Inc.
Keene, NH 03431

Introduction

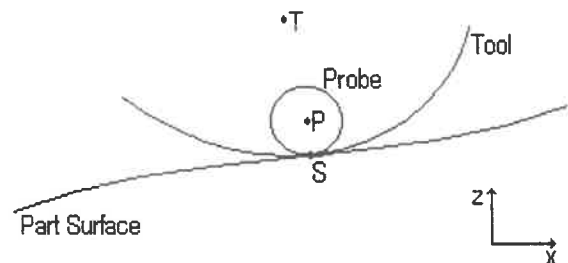
Modern methods of computer-assisted material processing, like diamond turning lathes and precision grinding machines [1], provide part surfaces with high-quality finish and of sub-micron form accuracy. Metrology control of these surfaces is a complicated and time consuming part of the manufacturing process, and is generally performed "off-line" on expensive stand-alone profilometry or interferometry devices [2].

The precision and repeatability of modern ultra-precise material processing machines are so high that in most cases the part form error originates not from machine inaccuracy but from the cutting process peculiarities, errors of tool path programming or from imperfect tool form. These errors can be investigated by using the cutting machine itself as a measuring device. This paper presents an on-machine technique for form error measurement and in-process corrective compensation.

System description

The system utilizes an ultra-precise air bearing Linear Variable Differential Transformer (LVDT) probe, mounted on a kinematic holder near the cutting tool. The kinematic holder allows the LVDT probe to be removed during the cutting process and subsequently replaced for measurement. The working linear range of the probe is $\pm 40\mu\text{m}$ with 10nm resolution. The probe collects data by traversing the machined part surface with its ruby ball tip. All special hardware and software for the probing and tool path correction resides on the machine's controller and forms a system called *Ultracomp*TM.

The input of the procedure is two blocks of data: initial tool path program, i.e. set of coordinates of tool origin (T on the scheme) along the tool path; and a set of coordinates of corresponding surface points S being processed (for each T_i point there is some S_i point)



Knowing the probe ball geometry, the *Ultracomp*TM system estimates the position of the probe ball (P on the scheme) for contacting every surface point S and generates a theoretical probe path for each specific surface. The LVDT probe follows that theoretical path and measures its real (x;z)-position.

The measured data is smoothed and interpolated to present a measured probe position for the same x-coordinates. The next step is to make the measured and theoretical probe paths correspond. If the position of the probing ball in the machine coordinate system is known with some error, the measured data is shifted with respect to theoretical, and their "difference graph" exhibits some slope. The optimization procedures shift the measured data to remove all x- and z-offsets. The difference of the z-coordinates of the measured and theoretical data represents the first order of part form error. Finally, the form error graph is displayed on the controller screen.

If one assumes that the part form errors are produced only by the errors of tool path, the subtraction of those errors from tool z-coordinates in initial NC-program should compensate the part program and correct form errors.

A special calibration procedure has been developed for recording probe geometry. Detailed mapping of the probe's ruby tip can be done in minutes by probing a master spherical artifact (traceable to NIST) of known radius. This procedure is similar to the measurement of a machined part: a theoretical probe path along the artifact is calculated based on the nominal ruby tip radius. The LVDT collects data, this data is compared with a theoretical one and form error data is generated. An iterative process estimates the best sphere fit radius for the tip. These radius and shape deviations of the probe ball from the sphere are stored in a special calibration file that is used for probe path generation. Non-spherical irregularities of ruby balls used in *Ultracomp*TM were found to be less than 80nm.

Discussion

Examples of part generation, measurement and correction cycle are shown in figs.1-2 together with comparison part interferograms. It has been demonstrated that the on-machine metrology / correction system allows a decrease of overall part form error from several microns to submicron level within a single iteration.

There are three ranges of surface error, caused by different physical reasons:

- small-range surface errors (noise, vibration, resonances, etc.), that are non-repeatable and affect mostly surface roughness;
- medium-range errors (tool defects, e.g. waviness) bring repeatable local form defects;
- long-range form errors (e.g. tool form differs from theoretical one).

The *Ultracomp*TM on-machine metrology system has been proven to detect all three given types of surface errors. It displays the form error graph and allows the technician to study the various parameters of the workpiece without removing it from the machine. However, the most effective application of the system is the on-machine correction of the repeatable medium and long range errors. These errors may be caused by tool waviness, positive or negative rakes, cylindrical clearances, etc.

Methods of surface measurement, data processing and tool path correction described above are currently being used both for diamond turning and periphery grinding. However, grinding often produces its own specific form errors.

Form errors unique to grinding include:

- 1) A sharp negative part form error (hole) in the center that originates from the fact that central area of the part has smaller longitudinal speed and the grinding wheel forces continually vary;
- 2) A form error produced by the error in grinding wheel radius estimation.

Since an error of the first type strongly depends on various grinding process parameters (material, wheel properties, cutting speed, speed of spindle rotation, etc.), it can be overcome empirically by making a trial cut, measuring the part and generating a new path that should compensate for the error. The measurement / compensation technique described here has proved to be a very effective method for removing the center hole error.

Wheel radius problems are very common in grinding. Since the wear of the grinding wheel (i.e. change of its curvature radius) occurs quickly, the precise tool geometry during tool path generation is very difficult to consider.

It can be shown that for the pure spherical part, error Δ in tool radius produces the same form error $|\Delta|$ in part curvature (same sign for a concave part and opposite sign for a convex part). In many cases of aspheric parts it can be assumed that the part profile function can be presented as the sum of aspherical and spherical components and their ratio can be considered $\ll 1$ for most of the part. If a best sphere fit radius for the measured part is known, then the data can "redistribute" form error according the following scheme:

$$\begin{aligned} \text{"Real part profile"} &= \text{"Design sphere"} + \text{"Design Asphere"} + \text{"Form error"} = \\ &= \text{"New best sphere fit"} + \text{"Design Asphere"} + \text{"New form error"} \quad (1). \end{aligned}$$

Therefore the deviation of the best fit sphere radius from the design sphere radius gives the estimation of the grinding wheel wear. The *Ultracomp*TM system performs this analysis which helps an operator estimate the error of part global curvature and the wear of the grinding wheel or diamond tool.

There are cases in optical design when tolerances of the radius of curvature of spherical components are less severe than for aspheric components. If in that case the radius of the "new best fit sphere" (1) satisfies tolerance requirements, we can use the "new form error" for tool path correction. Since that error can be substantially less than the full form error, that simplifies the compensation process.

Summary

Method of on-machine part form and surface measurement and compensation implemented in the *Ultracomp*TM system have proved to be equally suitable for both single point diamond turning and precision grinding. The system provides simple and fast metrological control of machined ultra-precise axially symmetric surfaces. Moreover, it can substantially reduce or eliminate repeatable form errors and improve surface quality.

References

1. Alexander H. Slocum, "Precision Machine Design", Prentice Hall, 1992
2. Daniel Malacara "Optical Shop Testing" John Wiley and Sons, 2nd ed., 1992

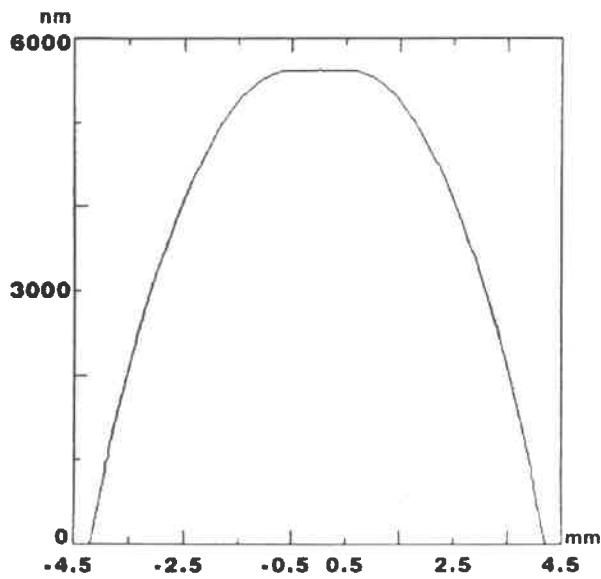


Fig. 1 The form error graph (left) of the diamond turned brass convex part (8mm radius of curvature) measured by the on-machine metrology system and the interferogram of the same part (right) for comparison. The on-machine system detects and eventually corrects not only the deviations from the spherical form seen on the interferogram, but also the general form error (error in part curvature) that is more than $5\mu\text{m}$ in this example.

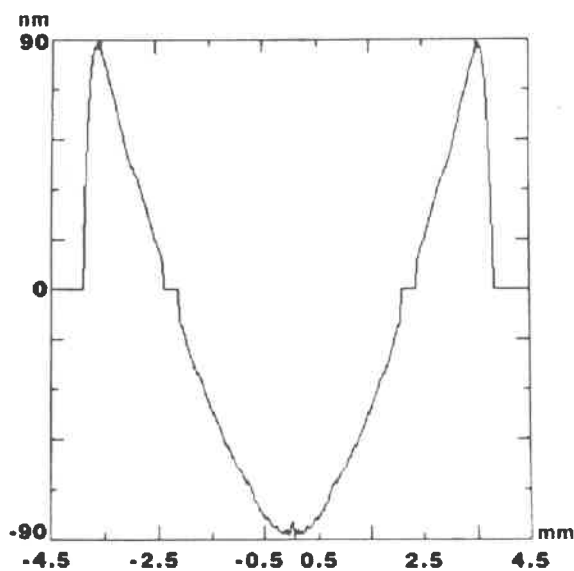


Fig. 2 The form error graph (left) and the interferogram (right) of the same part from fig.1 after the on-machine measurement and tool path correction have been performed. Total residual form error of the part is 180nm.

Development of Non-contact Measuring Machine for 3D Micro Pattern in Semiconductor Wafer using a New Optical Probe Implementing High Performance Optical Window

Heui Jae Pahk, Woo Jung Ahn

Dept. of Mechanical Design, Seoul National University, Seoul, Korea

Sung Wook Cho

Halla Optical Engineering, Seoul, Korea

1. Introduction

In recent years, quantitative measurement of 3 dimensional surface pattern has been increasingly applied in many science and engineering fields, and there have been several methods for the surface pattern measurement : Stylus based method, auto-focus based method, interferometric method, SEM(Scanning Electronic Microscope), and SPM(Scanning Probe Microscope), etc. In the case of semiconductor wafer, 3 dimensional pattern characteristics such as depth profile of chemically processed IC pattern was to be measured for the quality control of wafer manufacturing processes. Traditionally, SEM was used partly, but it caused inconvenience in wafer destruction and vacuum environment.

In this paper, a new type of non-contact/nondestructive measuring machine has been developed, comprising a new optical probe, precision stages, and controllers. The new optical probe has been designed and implemented around light illumination, optical lenses, optical window, and beam splitter.

Efficient image processing algorithm has been developed in order to analyze the 3D micro pattern. The developed system has been successfully applied to measure the wafer surface, giving few nanometer vertical resolution.

2. Theoretical Backgrounds for the Optical Probe

2.1. Blur Image

The new optical probe, which is a key feature of the developed measuring machine, has been newly developed. Fig.1 shows the basic principle of the developed optical probe for the height measurement.

Let I, II are the objects having δ height difference. When the objects I, II are transformed to image through the optics, the image A_1 , A_2 are formed at each focal plane, whose distances are b_1, b_2 from the lens, respectively.

When a CCD plane is located at the focal plane of the object I (b_1 distance), a sharp image of size A_1 and a blurred image of size R are formed at the b_1 focal plane. The blurred image is traditionally ignored or avoided, but it is fully used for the height measurement in this research.

From the law of geometric optics, the following two equations are derived.

$$\frac{1}{a} + \frac{1}{b_1} = \frac{1}{f} \quad (\text{Object I}) \quad (1)$$

$$\frac{1}{a - \delta} + \frac{1}{b_2} = \frac{1}{f} \quad (\text{Object II}) \quad (2)$$

, where f is the focal length of the lens.

From the geometric similarity of triangles,

$$a : b_1 = S : A_1 \quad (3)$$

$$a - \delta : b_2 = S : A_2 \quad (4)$$

where S is the size of object.

From eq(3),(4), the relationship between the height difference(δ) and the blur image size(R) can be constructed. That is,

$$\delta = a - \frac{ab_2}{b_1 \frac{A_2}{A_1}} = a - \frac{ab_2 A_1}{b_1 (A_1 + R)} \quad (5)$$

where a, b_1, b_2 are the distances from the lens, and A_1 is the image size of object I, and R is the size of the blur image.

Therefore when the sizes of object image(A) and blur image(R) are measured, the height data(δ) can be calculated from eq(5), if the coefficients are evaluated by the least squares technique.

2.2. Calibration Algorithm using Least Squares Technique

In order to obtain the accurate height data, δ , the calibration procedures are followed.

Prior to calibration, eq(5) can be arranged in simpler form, that is,

$$\delta = C_1 - \frac{C_2}{k} \quad (6)$$

where C_1 equals to a , C_2 is ab_2/b_1 , and k is the ratio of image size A_2/A_1 .

Based on measured data set (δ_i, k_i) for a master specimen, eq(6) can be expressed as follows.

$$k_i \delta_i = C_1 k_i - C_2 \quad (7)$$

For N data sets, the sum of squares of error, E , is

$$E = \sum_{i=1}^N (k_i \delta_i - C_1 k_i + C_2)^2 \quad (8)$$

Applying variational principle,

$$\frac{\partial E}{\partial C_1} = 2 \sum_{i=1}^N (k_i \delta_i - C_1 k_i + C_2)(-k_i) = 0$$

$$\frac{\partial E}{\partial C_2} = 2 \sum_{i=1}^N (k_i \delta_i - C_1 k_i + C_2)(1) = 0$$

Thus

$$C_1 = \frac{\begin{vmatrix} \sum k_i^2 \delta_i & -\sum k_i \\ -\sum k_i \delta_i & N \end{vmatrix}}{\begin{vmatrix} \sum k_i^2 & -\sum k_i \\ -\sum k_i & N \end{vmatrix}} = \frac{N \sum k_i^2 \delta_i - \sum k_i \sum k_i \delta_i}{N \sum k_i^2 - (\sum k_i)^2}$$

$$C_2 = \frac{\begin{vmatrix} \sum k_i^2 & \sum k_i^2 \delta_i \\ -\sum k_i & -\sum k_i \delta_i \end{vmatrix}}{\begin{vmatrix} \sum k_i^2 & -\sum k_i \\ -\sum k_i & N \end{vmatrix}} = \frac{-\sum k_i \delta_i \sum k_i^2 \delta_i + \sum k_i^2 \delta_i \sum k_i}{N \sum k_i^2 - (\sum k_i)^2} \quad (9)$$

Therefore coefficients C_1, C_2 are evaluated using the least squares technique, completing relationship between the height data, δ , and the ratio of image size, $(A_1+R)/A_1$.

2.3. Optical Window for Light Slit and Image Processing Algorithm

The optical window is inserted between the lenses and the splitter for making light slit and windowing the object size. Fig.2 shows the measured image of several step heights on a wafer surface. In the Fig.2, the sharp image is the image of reference step height, and the blur images are the images of step heights to be measured. Using the principles of the blur image, the step height data can be calculated along the vertical line. The object images are captured by a CCD camera associated with

magnifying optics system. The captured image is then processed to give the height data using image processing algorithm developed. The measured step height data are shown on the left of the figure. Therefore step height measurement has been implemented using the optical window and associated optical principles. Also, the feature of line scanning has been implemented.

3. Measurement System

The developed measuring machine consists of the optical part with illumination, mechanical part for motorized wide range scanning, and software part for operating measurement system.

Fig.3 shows the block diagram of optical and illumination part. The image frame grabber is implemented for capturing CCD camera image, and a light controller is also used for the illumination control. The whole optical image processing and illumination control are performed under Windows based image analysis and light control software developed.

For scanning wide range of wafer surface, a precision X-Y stage has been designed using air bearing, linear motor, optical scale, granite surface plate. Z-W-Theta stage also has been designed for the wafer handling and focus adjustment. Fig.4 shows the block diagram of mechanical motion control, where brushless linear motor, stepping motor, DC motor controllers are implemented. The whole motion control is operated by the Windows based motion control software developed. Thus it is possible to select specific measurement location within wafer surface by the menu driven software developed.

Fig.5 shows the non-contact measuring machine developed.

4. Practical Application to Wafer Surface

The developed measuring machine has been practically applied to DRAM wafer surface. The scanning area is chosen as 20um x 20um, and the optical probe line-scans the wafer surface at 0.1 um interval in the horizontal direction. Fig.6 shows the 3D plot of the measured wafer surface, giving height properties such as flatness. Three dimensional roughness parameters have been successfully evaluated by the developed measurement software, based on the optical measurement data. The measurement resolution has been achieved around few nanometers.

5. Conclusion

- (1) A new measuring machine for 3D micro pattern of wafer surface has been developed, giving nondestructive measurement feature with less than 10 nanometer resolution.
- (2) A new optical probe has been designed and implemented using the blur image of non-focused object. The feature of line-scanning capability of the optical probe has been found very much effective in practical measurement application.
- (3) For scanning wide range of wafer surface, precision XY stage, Z-W-Theta stage have been designed, and the measurement region can be easily selected within whole wafer surface by user-interactive software developed.
- (4) Variety of surface properties such as flatness, 3D roughness parameters can be evaluated by the measurement software developed, and is believed to replace SEM in many 3 dimensional pattern measurement, as the developed system does require neither vacuum environment nor specimen destruction

Reference

1. M. Subbarao, G.Surya, "Depth from Defocus: A Spatial Domain Approach", International Journal of Computer Vision, 13, 3, 271-294, 1994
2. M. Subbarao, Tae Choi, "Accurate Recovery of Three-Dimensional Shape from Image Focus, IEEE Trans. On Pattern Analysis and Machine Intelligence, Vol17, No.3, March 1995

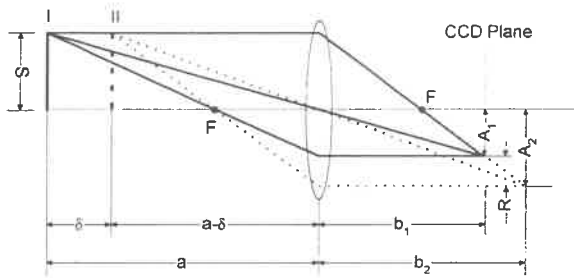


Fig.1 Basic Principle of the Optical Probe

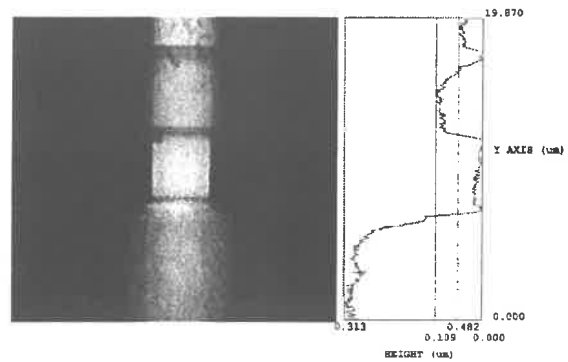


Fig.2 Image of Measured Area

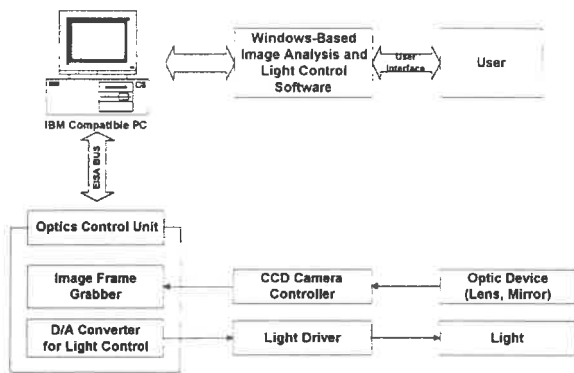


Fig.3 Block Diagram of optical and illumination part.

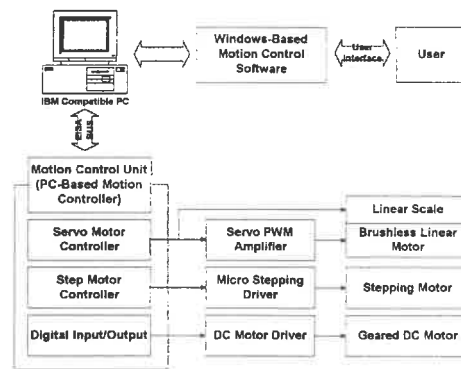


Fig.4 Block Diagram of Mechanical Motion Control



Fig.5 The Developed Non-contact Measuring Machine

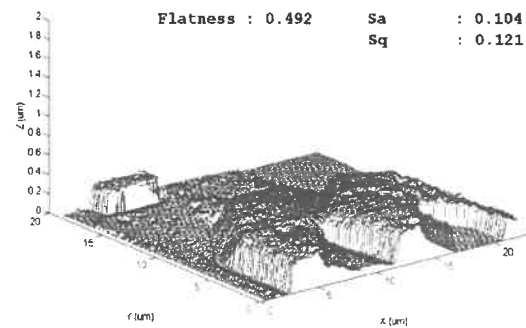


Fig.6 3D Plot of the Wafer Surface Measured

Development of a 2D probing system with nanometer resolution

Pril W.O., Struik K.G., Schellekens P.H.J.

Eindhoven university of technology

Eindhoven, the Netherlands

Abstract:

A newly developed 2D probing system for a small Coordinate Measuring Machine (CMM) with nanometer resolution and low probing force is presented. The instrument has been built and the first tests have been carried out, which will be presented here.

The probes measuring system is based on a Laser-Diode-Grating-Unit (LDGU). Used in CD-players, it is a cheap, mass produced component. The LDGU here measures two degrees of freedom of a mirror attached to the stylus: vertical translation and a tilt around one horizontal axis. The resolution has been measured to be 1 nm in vertical direction and 100 nm in one horizontal direction.

Special attention has been paid to the static and dynamic probing force to prevent indentations on the workpiece. The stylus with the attached mirror is kinematically suspended from the probe housing by three small leaf springs. The stiffness of the suspension is very low, allowing a static probing force of 1 μN . The dynamic measuring force when probing the workpiece can be orders higher than the static force, especially when the mass directly connected to the stylus is high. A very low moving mass of 4 mg limits the dynamic force to some mN or less, depending on, among others, contact velocities and probe ball diameter. This is small enough to prevent even very soft materials from being damaged. The possibility of measuring elastic samples, e.g., small plastic components, is another significant advantage of this low probing force.

Several precautions have been taken to prevent thermal drift. The probe housing is made from invar. The heat (80 mWatt) generated by the LDGU is conducted out of the construction by a thermal path that is to a large extent separated from the mechanical path.

Introduction

As miniaturization goes on, tolerances on dimensions become smaller and smaller. Often a Coordinate Measuring Machine (CMM) is used to check these dimensions because of its versatility and ease of use. However, for the most accurate parts, made e.g. on precision turning machines, two problems with the use of commercial CMM's arise:

- The CMM inaccuracy is 0.5 μm or more, which is comparable to the accuracy that can be reached with nowadays precision manufacturing techniques.
- Probing forces of normal CMM probes are too high for the materials that are often used for precision products. Measuring forces of 0.1 N or more are common. At the moment of probing, however, the (dynamic) forces are mostly much higher¹. These forces may damage the surface of certain materials, such as aluminium, brass and even steel.

In our group two CMM's are being developed, one with an inaccuracy below 0.1 μm in a measuring volume of 1 dm^3 ², the other especially for scanning, with a measuring volume of several cm^3 and a very fast z-axis with an accuracy in the nm region. Both CMM's require a newly developed probing system with high accuracy and low probing forces. A 2D probing system fulfilling these demands is described here. It is mainly intended for the scanning CMM, where two dimensions will satisfy in most cases. The second dimension is advantageous compared with 1D probes, because they may have problems detecting steep edges. In the meantime valuable experience can be gained for developing a full 3D version.

Working principle

In figure 1 the working principle of the probe system is shown. The stylus (S) is suspended elastically (c) to the probe house (details will follow). An optical measuring system was chosen because it is contactless and only a small mirror (M) has to be attached to the stylus. This is important because the smaller the moving mass of the probe, the smaller the probing forces will be as will be shown in the next paragraph. The measuring system uses a Laser Diode Grating Unit (LDGU) used in commercial CD-players³. Due to mass production this is a cheap but nevertheless accurate measuring tool. It contains a laserdiode (L), four photodiodes (D1...D4), and a grating (G). The light from the laser is focused by two lenses (L1 and L2) on the mirror and then reflected back into the LDGU. Passing through the grating the first order beam is split in two half beams which are imaged on D1 till D4. Depending on the imaging of the focal point of the half beams in front of or behind the diodes, D1 and D4 or D2 and D3 are illuminated. The Focal Error Signal (FES) is defined as³:

$$FES = \frac{D1 + D4 - (D2 + D3)}{D1 + D2 + D3 + D4} \quad (1)$$

where D1..D4 are the signals of the photodiodes. In the transitional zone the sensitivity for z-displacement of the mirror is very high. Besides z-position, the LDGU can also measure tilt of the mirror in one direction (RES: Radial Error Signal)³:

$$RES = \frac{D1 + D2 - D3 - D4}{D1 + D2 + D3 + D4} \quad (2)$$

The focal lengths of L1 and L2 determine the sensitivities of the system:

$$S_{RES} \propto \frac{f_2}{f_1} \quad S_{FES} \propto \left(\frac{f_1}{f_2} \right)^2 \quad (3)$$

Two lenses with a small focal length of 9 mm were chosen, leading to high S_{FES} in comparison with S_{RES} and a compact design. The photo currents of the photodiodes are about a microampere which must be measured with an inaccuracy of a few nanoamperes. Therefore the wiring between the diodes and the preamplifiers may not be longer than a few centimeter, so preamplifiers were integrated in the probe system. Micropower opamps, consuming no more than about one mWatt each, are used to limit thermal drift. An SMD version of the IC makes them fit in the very limited space available in the probe system. Before the mechanical details will be discussed, first some remarks on probing forces will be made, because two important design parameters will be derived from this.

Probing forces

As mentioned earlier, the probing force is a big problem when using CMM's in precision metrology. In order to estimate how low the probing force should be to prevent plastic deformation, a model of the probing procedure was made, part of which is described by van Vliet¹. It is also used to deduce design parameters. Three forces, originating from different events, can be distinguished:

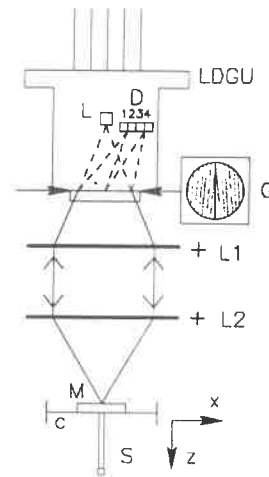


figure 1: working principle

- Collision force F_{col} . This force acts at the moment when probing the sample. The moving parts of the probe, i.e. stylus and moving part of the measuring system, have to be decelerated, which leads to forces between probe and sample. During this collision the probe may bounce once or several times.
- Overtravel force F_{ovt} . After receiving a stop signal from the probe, the CMM moves on a certain distance, called overtravel, which stretches the springs of the probe suspension leading to an elastic force between sample and probe. Because the maximum of F_{ovt} occurs when F_{col} has decreased to almost zero, both forces can be compared with the admissible force F_Y separately.
- Measuring force. This force works between sample and probe when the data point is taken. Its magnitude is usually adjustable. Although this is the force specified for most probe systems, it is not really important when considering damaging of the workpiece, because it is generally much smaller than the previously mentioned dynamic probing forces.

Both F_{col} and F_Y can be calculated using Hertzian spherical contact theory, described in e.g. Szabó⁴ and van Vliet¹. Derivation in other directions is not mentioned here because the forces in z-direction are most critical in our situation.

$$F_Y = 2l \frac{R^2 Y^3}{E^{*2}} \quad [N] \quad (4)$$

$$\frac{l}{E^*} = \frac{l - \nu_1^2}{E_1} + \frac{l - \nu_2^2}{E_2} \quad [m^2/N]$$

$$F_{col} = \left(\frac{5mv^2}{4\alpha} \right)^{\frac{3}{5}} [N], \quad \alpha = \left(\frac{9}{16E^{*2}R} \right)^{\frac{1}{3}} \left[m/N^{\frac{2}{3}} \right] \quad (5)$$

As can be seen in (5) F_{col} and F_Y are dependent from a number of parameters:

E_1, E_2	Young's moduli of probe ball and object material
ν_1, ν_2	Poisson moduli
Y	Yield strength of the workpiece (provided that the yield strength of the probe ball is higher). In practical situations comparable to $\sigma_{0.2}$.
R	Probe ball radius
v	probing speed
m	Moving mass of the probe

Because of the expected size of the workpiece it should be possible to measure with small probe ball diameters (e.g. 0.3 mm), which leads to an admissible probing force of one to several mN. For practical reasons v should be at least 1 mm/s, which limits the moving mass to several milligrams at most for reasonable object materials.

Assuming an overtravel of 5 μ m, being the design value for our dm³ CMM at a probing speed of 1 mm/sec, it follows from (4) that $c_z < 0.1 \text{ N/mm}$, where c_z is the stiffness by which the moving part of the probe is connected to the probe house. In spite of this small stiffness the eigenfrequency of the probe is still very high, namely 800 Hz.

Elastic Suspension

The stylus is suspended elastically to the probe house by three leaf springs (figure 2) to enable sufficient overtravel of the CMM. The stylus is connected to them by an intermediate part. The leaf springs give the stylus three degrees of freedom: translation in z-direction, and rotation along x- and y-axis. The rotations give rise to pseudo translations at the probe ball position in y and x respectively, so it can be moved in three directions. No other intermediates are needed for this suspension so that the moving mass can be kept small. Another advantage is the ease of manufacturing: because the leaf springs are all in one plane it can easily be etched from sheetmetal. Using elastic theory the stiffness in all directions can be calculated, providing that the displacements are small⁵:

$$c_z = 3 \frac{Ewt^3}{l^3}$$

$$c_{xy} = \frac{3Ewt^3}{ll_s^2} \left(\frac{1}{2(1+\nu)} \alpha_t - \frac{y}{l} + \frac{1}{12} + \frac{y^2}{l^2} + \frac{x^2}{l^2} \right) \quad (6)$$

where E and ν are Young's and Poisson moduli, and t the thickness of the leaf springs. l_s is the stylus length and α_t is a torsional 'constant' depending on t/w . It approaches $1/3$ for small t/w . Note that the stiffness is independent from the direction of probing in the xy-plane. In order to get F_{ov} below F_Y , c_z and c_{xy} should both be smaller than 100 N/m, which is obtained by using 20 μm thick spring steel for the leaf springs.

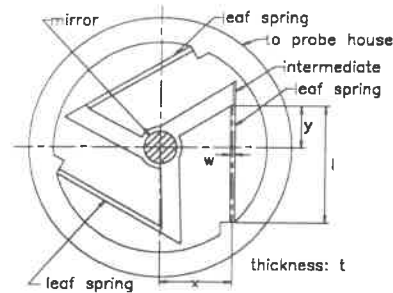


figure 2: elastic suspension

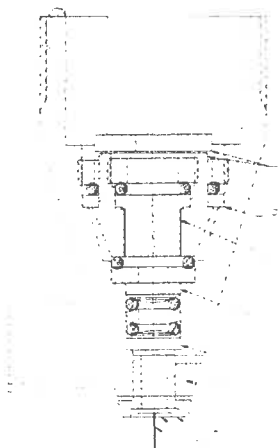


figure 3: construction

Mechanical design

LDGU (1), lenses (2,3) and a subassembly with leaf springs (4), mirror (5), and stylus (6) have to be positioned with nanometer stability (the numbers refer to figure 3). This is extremely difficult because the laserdiode in the LDGU dissipates about 80 mWatt. Therefore most parts were made from invar. To conduct the heat out of the construction an aluminum thermal path (7) was included, insulated by plastic (8) and air from the rest of the structure. Although the LDGU is fixed by the aluminum, no thermal length is bridged by it. Outside of the construction the heat should be transferred to the air by heatsinks, either passively, or actively by using Peltier elements.

An adjusting mechanism (9) was added to fix the mirror within the range of the measuring system. A lot of attention was paid to the outer dimensions: The probe system should be as narrow as possible so the tip can be seen from a reasonable angle.

Measurements

Measurements in z-direction were done, using a calibrated laser interferometer as a reference. A piezo actuator was used to move the stylus over a 3 μm distance. Due to the use of a C-frame support it is impossible to evaluate drift, because the expected considerable drift in the support cannot be distinguished from the drift in the probing system.

A better setup is being developed, but it's not yet available at the time of writing. The resolution, limited by the noise, is measured to be less than 1 nm. A 30 nm top-top non-linear effect is superimposed on the signal (figure 4), most probably caused by back scattering into the laserdiode or interference of the light on the photodiodes with stray light. For the time being the inaccuracy is limited to about 40 nm because of the non-linearity. Further research has to be done to find the origin of the effect and to correct it. It is expected that this will improve the accuracy to 10 nm (expected value).

The same experiment was repeated in x direction. The resolution in this direction is measured to be 100 nm.

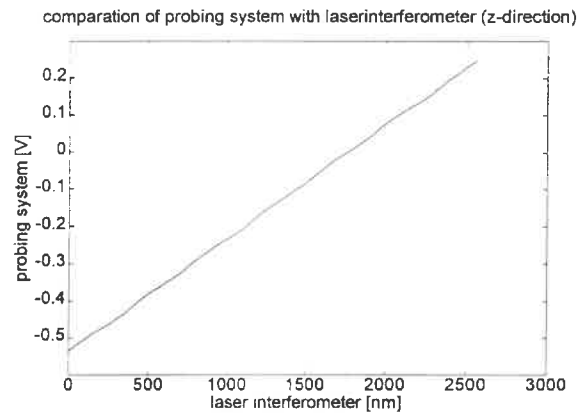


figure 4: comparison of probing system with laser interferometer

Conclusions

A 2D probing system was designed which combines high resolution with low probing force. The resolution has been measured to be 1 nm in z-direction and 100 nm in x-direction. Low masses in the design and a suspension with a low stiffness limit the dynamic probing force to several mN or less, depending on a number of parameters. This force is low enough to prevent damaging of the workpiece and influencing of the measurement. An interference effect seems to be superimposed on the FES signal so that the inaccuracy is limited to 40 nm. Further research has to be done to correct this and improve the accuracy to 10 nm.

¹ W.P. van Vliet, P.H. Schellekens; Annals of the CIRP **45**, 483-487, 1996.

² M.M.P.A. Vemeulen, P.C.J.N. Rosielle, P.H.J. Schellekens. Proceedings from ASPE Annual Meeting **10**, 133-137, 1994.

³ Sharp Technical information, model LTOH10M, (1988).

⁴ I. Szabó; Höhere Technische Mechanik (in German). Springer-Verlag Berlin Heidelberg New York, 1977.

⁵ W.O. Pril: Ontwikkeling van een 3d-CMM-tastsysteem met nanometernauwkeurigheid (in Dutch); Thesis post-graduate course Mechatronic design 1995, ISBN 90-5282-479-7.

The Calculation of CMM Measurement Uncertainty via The Method of Simulation by Constraints

S.D. Phillips, B. Borchardt, D. Sawyer, W.T. Estler, D. Ward, K. Eberhardt, M.S. Levenson, M. McClain,
National Institute of Standards and Technology
B. Melvin, ICAMP Inc., T. Hopp, ZigZag Inc., Y. Shen, G. W. University

Introduction and Background: The calculation of task-specific measurement uncertainty when using coordinate measuring machines (CMMs) is an important and challenging task. By task-specific measurement uncertainty we mean the uncertainty of a particular feature inspected with a specific inspection plan and stated as prescribed by the *ISO Guide to the Expression of Uncertainty in Measurement*. The complete problem involves the assessment and propagation of numerous uncertainty sources including imperfect parts and extrinsic factors (see Fig. 1). One aspect of this problem which has attracted considerable attention is the assessment and propagation of uncertainties related to the measurement instrumentation, e.g. the CMM. Even this limited problem, (which assumes perfect parts, no extrinsic factors, and accurate software) represents a daunting task because CMMs can inspect an almost limitless variety of features, in a wide variety of locations and orientations on the CMM, using an infinite variety of sampling strategies (the number and locations of the probing points).

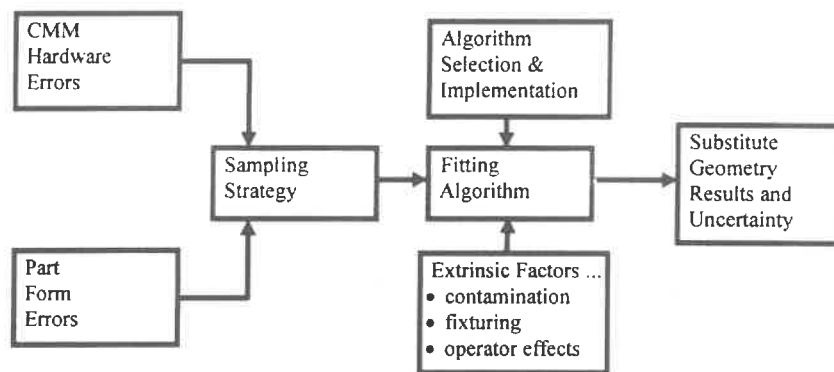


Figure 1. Schematic outlining the various factors affecting CMM measurements.

Historically, performance specifications embodied in national or international standards (which were designed to promote commerce) have been unable to provide useful estimations for task specific uncertainty, unless the part under inspection happened to be nearly identical to the calibrated artifact used in the performance evaluation. This has led to the development of the virtual instrument concept where software can propagate errors into the feature measurement under consideration. For CMMs this has involved the assessment of the individual rigid body parametric errors, i.e. the translational and rotational errors of each axis, and their propagation through the kinematic equations describing the CMM structure, to yield point coordinate errors. Point coordinates which contain these and other CMM hardware errors (e.g. due to the probe) are then run through the appropriate geometrical fitting algorithm to yield the error in the final feature.

Full Parametric Simulation (FPS): Recently, this process has been expedited (for small CMMs) by the use of calibrated ball plates measured in different positions and orientations from which one can extract the individual rigid body parametric errors. If all the parametric errors were known exactly and rigid body kinematic equations exactly described the CMM, then their propagation would determine the measurement error (the measured value minus the true value). However, since there is uncertainty in the calibration of the ball plate, and the kinematic description is only approximate, this calculation estimates the measurement bias of the feature due to the rigid body errors of the CMM. Recognizing that the parametric errors are not exactly determined, this uncertainty is assessed by simulating small perturbations about the measured mean parametric errors using many different measurement scenarios each with a slightly different parametric error description. The collection of calculated feature results (each of which will be slightly different reflecting a slightly different parametric state) is descriptive of the uncertainty in the feature due to the imperfect ball plate and the reproducibility of the CMM. This uncertainty can be combined with the estimated bias to yield the uncertainty in the feature due to the CMM; any additional sources of uncertainty such as part form errors or extrinsic effects must then be combined to produce a final uncertainty statement for the feature. This procedure, which in many situations can accurately describe the CMM's contribution

to the measurement uncertainty, has become known as the “virtual CMM”; however, we prefer the somewhat more descriptive name of “Full Parametric Simulation” (FPS). Its essential feature is that it requires a complete description of the CMM parametric errors.

Simulation by Constraints (SBC): We will now outline a generalization of this technique, of which FPS is a special case. Consider a three axis CMM. We call the “parametric state of the CMM” the particular values of the various rigid body parametric errors of the CMM. In general this unknown vector will include 18 functions (pitch, roll, yaw, etc.) and three scalars (squareness values). Mathematically, we can think of the parametric state of the CMM as being represented by a particular point in a large infinite dimensional state space. If we have no prior performance information about the CMM then we are completely ignorant of where in the state space the CMM actually lies; hence it could be said that our uncertainty about the parametric state fills all of state space. Suppose we learn that this CMM has measured a calibrated step gauge oriented along its X axis and that the largest error is less than $10\text{ }\mu\text{m}$. From this information we can reasonably state that it is unlikely the X axis scale error $\gg 10\text{ }\mu\text{m}$. Although it is possible that the various parametric errors combine with the X scale error in a manner such that they largely cancel each other out, this hypothesis becomes increasingly unlikely if we learn that additional step gauge measurements (oriented along the X axis but translated in the Z direction) continue to have errors less than $10\text{ }\mu\text{m}$. Hence with a few measurements we can reasonably state that the X axis parametric scale error is bounded by $\eta \cdot 10\text{ }\mu\text{m}$, where η is greater than but on the order of unity. In terms of the state space, we can say that the subspace representing the X axis scale error is bounded. However, since there are no bounds on the other parametric errors, the region of state space which may contain the state of the CMM (and still be consistent with the known information) is still unbounded. Thus, in this example, our measurements form a constraint sufficient to bound one of the parametric errors but insufficient to provide a topologically bounded region in state space. Our goal is to obtain measurements which result in a topologically bounded region in state space. We call such a set of measurements a Bounding Measurement Set (BMS). Any calibration of the CMM must yield a BMS as a result of the calibration procedure, otherwise some error can have an arbitrarily large value. The points inside the bounded region of state space are known as “virtual states” because any one of them could be the “true” state of the CMM. The virtual states represent our ignorance, i.e. our uncertainty, about the exact state of the CMM. The boundaries of the virtual states represent our knowledge about the CMM, i.e. they summarize our information about the CMM, and all virtual states inside the boundaries are consistent with this information.

As a conceptual illustration of SBC, Fig. 2 depicts an idealized instrument which possesses only two error sources denoted α and β , both of which are scalars, e.g. these might be the relative scale error between two axes and a squareness error. In this idealized example, the state space is two dimensional and specifying the values of the two error sources uniquely locates the state of the CMM. Shown in the figure is the estimated state of the CMM (based on a full set of parametric measurements), the associated bounded region determined by FPS (this includes the true - but unknown -- state of the CMM). Also shown are two bounded regions derived from a BMS, i.e. multiple topologically bounded regions of state space may be derived from the BMS. (Fig. 2 illustrates a general case, however, in many situations there will be a symmetry of the bounded regions of virtual states permitted by BMS due to the symmetry in the kinematic equations and the BMS in terms of the error parameters α and β .)

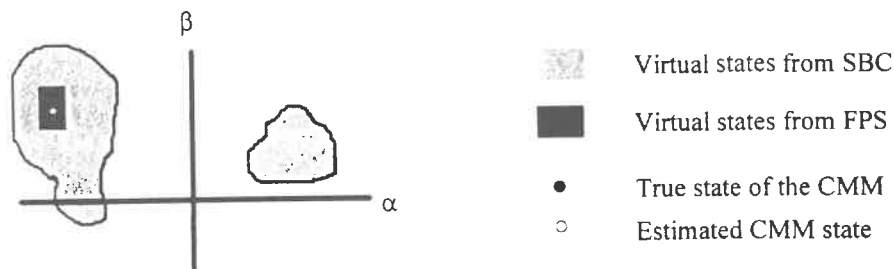


Figure 2. Schematic of a two dimensional state space

Central to the SBC method is that estimates of each parametric error are not required (as is the case for FPS). The only requirement is that sufficient measurements are taken to bound the parametric errors, i.e. a BMS. Hence SBC works with incomplete information about the state of the CMM whereas FPS uses a complete description of the CMM state. To illustrate the magnitude of this difference, a typical CMM might have each of its 18 functional parametric errors measured at a dozen points along each axis, plus three scalars (squareness values), resulting in

over 200 numbers to characterize its state. In contrast we will show that by using five numbers, i.e. five CMM performance evaluation test results, we can obtain very reasonable uncertainty estimates. Hence by using less than 5% of the amount of data needed for FPS, SBC can achieve useful results. The key to SBC is that the small amount of data utilized must be of high quality in the sense that it forms a good BMS.

SBC Using The ANSI/ASME B89.4.1 Standard: Ideally, the data used in the SBC method would be widely available to the user, as is the case for published performance specifications given by national or international standards. We will present the case of the US national standard which characterizes the geometric errors of a CMM, i.e. those involving the parametric errors, by five numbers: linear accuracy results for the X,Y, and Z axes, the volumetric performance, and the volumetric performance using an offset probe. With some assumptions (described later) these five values can form a BMS. Conceptually, SBC proceeds in the following manner: a random parametric state is selected from all of state space and the B89 performance tests are simulated using this state. If the simulated test results agree with the known five B89 values we label this a "good virtual state". If the simulated and known B89 results disagree the trial state is rejected and another random state is selected and tested. Continue this procedure until a large set containing several hundred good virtual states is obtained. This set represents a sampling of the topologically bounded regions in state space, and using these virtual states any part measurement under consideration can be simulated. The resulting collection of part feature results represents an estimate of the feature measurement uncertainty due to geometrical errors of the CMM. To test the uncertainty prediction of the SBC method we have measured a calibrated step gauge in 43 different positions and orientations throughout the workzone. In each position five different lengths were measured, and each length was measured three times. Figure 3 shows the SBC calculated expanded uncertainty ($k = 2$) and the actual measurement errors for the longest of the five lengths (this is 1000 mm for all positions except along the Z axis which is 700 mm), for a CMM with large uncorrected geometric errors. Figure 4 displays the same data as Fig. 3 but for a (parametric) error corrected CMM. For these two figures, as well as the eight similar figures (not shown) associated with the shorter length measurements, approximately 95% of the measurement errors are contained within the SBC calculated expanded uncertainty. We consider these results, based on a paucity of input information, to be quite reasonable.

Implementation Issues of SBC and the B89 BMS: Our implementation of the SBC method assumes that the parametric errors are analytic and slowly varying, e.g. the spatial frequency of the errors is restricted to less than 50 cycles/meter. Additionally, rather than picking prospective virtual states at random, we employ a reasonable initial guess and then use a gradient search method to locate good virtual states. (Some caution must be exercised to avoid a search that preferentially seeks only certain areas in state space, otherwise the search may miss disconnected regions of valid virtual states, see Fig 2). Using the B89 test results as a BMS also requires a few reasonable assumptions. Specifically the B89 tests are very weak at detecting straightness errors; consequently we employ a general kinematic relation between angular errors and straightness errors to bound this parametric error source. Similarly, only the linear term of the Z axis roll error (for a moving bridge machine) is constrained by the B89 tests; consequently we force the nonlinear terms to be smaller than the linear term in magnitude. This procedure can be somewhat justified both by experience and by the fact that the exact location in the workzone where the point-to-point Z roll test is performed is (in practice) determined by the user.

We have outlined the method of SBC for CMM uncertainty calculations. Using incomplete information about the CMM we are able to construct realistic point coordinate uncertainties which can, in turn, be used to calculate feature measurement uncertainty. Although feature measurement uncertainty is the desired objective, mathematically the challenge is to efficiently locate the good virtual states as these determine the point coordinate errors. The feature fitting algorithm can be considered simply a transformation from point coordinate errors to feature errors (see Fig. 1). The method is very general and can be applied to diverse systems, e.g. machine tools and laser trackers. Similarly, many different measurement schemes can be used for the BMS. The use of the B89.4.1 standard works well provided reasonable assumptions are made about straightness and Z roll errors. Finally, the standard could form a stronger constraint if measurements sensitive to straightness and the nonlinear component of Z roll were included and if the linear accuracy and volumetric performance values were stated as deviations from calibrated values instead of ranges. In the limit of taking more and more measurements the constraints become stronger. Ultimately only one set of parametric errors is consistent with the data, and we have reached the domain of FPS.

Acknowledgments: This work was funded in part by the Air Force CCG Program, NIST's Computational Metrology Project, and the Manufacturing Engineering Laboratory's NAMT Program.

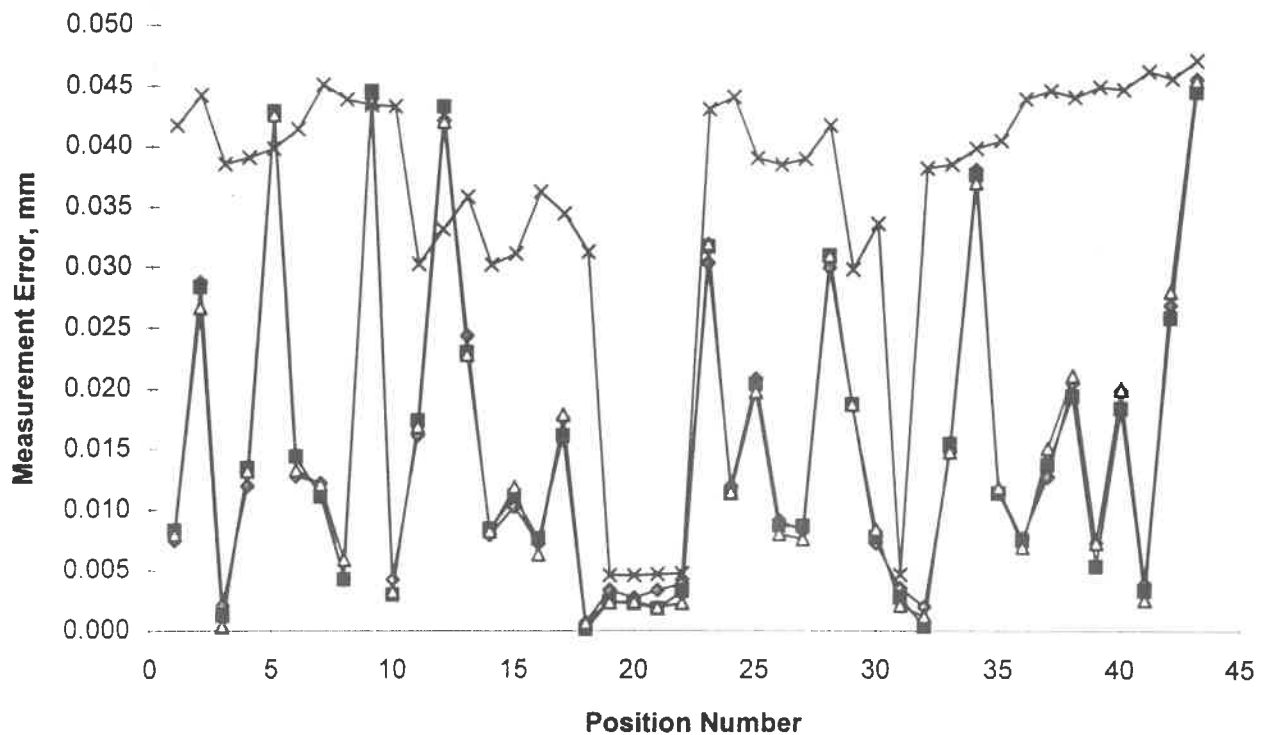


Figure 3. The calculated expanded uncertainty via the SBC method (shown as x) of 43 point-to-point length measurements of a step gauge and three sets of measured deviations from the calibrated value (■, ◆, △). The constraints for this CMM are: X, Y, Z linear accuracy are 13.8 μm , 18.3 μm , 3.4 μm respectively, volumetric performance = 39.1 μm , volumetric performance with an offset probe = 55.0 $\mu\text{m/m}$.

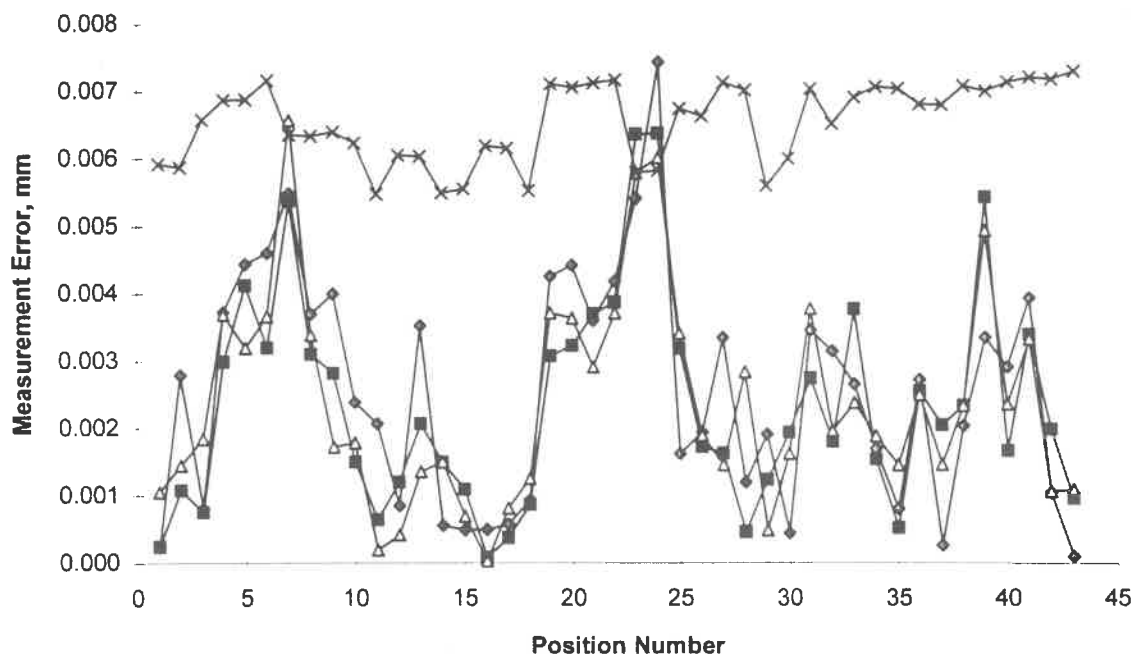


Figure 4. The calculated expanded uncertainty via the SBC method (shown as x) of 43 point-to-point length measurements of a step gauge and three sets of measured deviations from the calibrated value (■, ◆, △). The constraints for this CMM are: X, Y, Z linear accuracy are 3.0 μm , 2.1 μm , 2.5 μm respectively, volumetric performance = 7.2 μm , volumetric performance with an offset probe = 7.1 $\mu\text{m/m}$.

Thermal Responses of Coordinate Measuring Machines

C. Nickles, T. Charlton
Brown & Sharpe Manufacturing Company
North Kingstown, RI

Coordinate measuring machines are being used with increasing frequency in environments that are thermally hostile to accurate measurement. The driving force for this change is the great benefit of more closely coupling precision measurement to manufacturing processes by gauging on the shop floor. The chief obstacles for this change are (a) objects only have their correct size at 20 C, (b) thermally-induced measurement errors represent differential expansion between the workpiece and measuring instrument elements, and (c) uncertainties associated with thermal correction techniques. The conflict between the movement of CMMs to the shop floor and the resulting thermally induced errors has led to a variety of claimed solutions for thermal problems. These include using mathematical compensation techniques, thermally insensitive machine components, localized temperature control methods, and error-budget uncertainty estimation. Unfortunately, no well-defined objective criteria exist for characterizing either the success or suitability of any of these possible solutions, or even the real magnitude of the problem.

To improve the quantitative discussion of these issues, we report on the development of a standard stimulated-response test for thermal effects. Our approach evolved out of our extensive experimentation with the thermal responses of CMMs and subsequent attempts to explain the results to diverse audiences. The current development addresses three essential features of characterizing thermal responses: the temperature "input" profile, the measurement task definition, and the simplified presentation of the results.

Temperature Cycle Selection

To "stimulate" differential thermal responses in the CMM and workpiece, we apply a specific environmental temperature profile as a function of time. By selecting an appropriate profile our goal is to produce test results representative of shop floor errors. Having experimented with sinusoids, step functions, impulses and various triangular pulses, we finally prefer a repeated diurnal cycle shown as Figure 1. Workpiece measurement accuracy is more widely and routinely degraded by diurnal cycling than by any other thermal effect. Users and researchers alike can readily accept the reality of this proposed cycle, if not its consequences. Further, this profile is reasonably easy to produce as it only combines linear ramps and constant dwells over a limited temperature range. Note that the specific details of one diurnal cycle was selected to qualitatively mimic a Summer workshop (middle Northern latitude) with two work shifts, including warming due to machinery power-up and solar ascension, cooling due to solar descension, and finally cooling due to machinery power-down.

We emphasize that standardizing on a temperature profile is essential to make possible direct comparison of results from different machine designs with different workpiece materials while using different techniques to handle the temperature effects.

Measurement Task Selection

The second essential feature of our work is the selection of a simple artifact as a surrogate workpiece, and the specification of the measurement sequence to be used. As with the temperature profile, dual goals of standardizing on one workpiece geometry are to facilitate reproducibility and comparative studies and to produce representative shop-floor results. Expanding on the excellent experience with planar ball plates (ref: Kunzman), we favor an inexpensive modified ball plate shown in Figure 2. The spheres roughly span a 300 mm cube. Key modifications are the addition of a perpendicular spar and the addition of a centralized clamping feature.

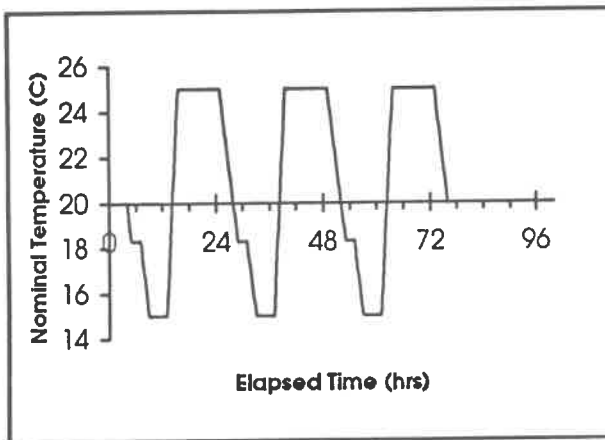


Figure 1. Standard Thermal Test Cycle

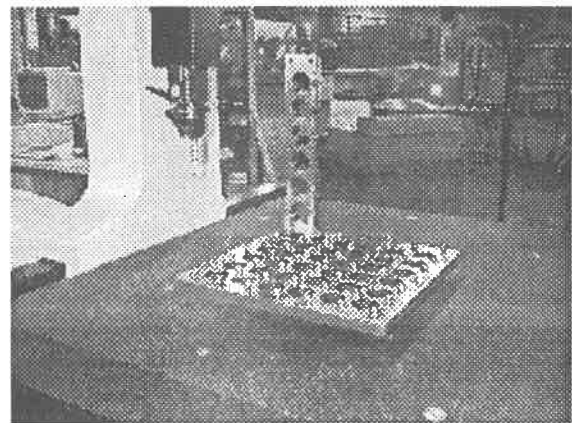


Figure 2. Thermal Test 3D Ball Plate

The addition of the spar allows measurement of the plate to incorporate some full three dimensional measurements in every orientation of the plate on the machine. This provides a sensitivity to changes in all of the machine axis squarenesses, which is a frequently unrecognized part of the thermal errors. Originally designed with kinematic location, our plate design now uses a single point clamping method. This change was made to reduce the influence of variable clamping forces and the resulting variable effects of different coefficients of expansion between the workpiece and the workpiece mounting surface. Such variability of thermal response due to fixturing of workpieces is extremely significant in practice. Unfortunately, no mounting method is reasonably representative of fixturing induced thermal response problems. Thus, it seems best to simply minimize these effects for a standard performance test, while emphasizing its importance.

To produce representative results for the application of CMMs it is important to report and compare functional results for workpiece materials with significantly different coefficients of expansion. Since the most significant engineering materials inspected on CMMs are steel and aluminum, plates of both of these materials were fabricated and used. Fortunately, these two materials cover a relatively wide range of expansion coefficients.

In applying this ball plate for thermal testing, we repeatedly measure the center coordinates of a fixed sequence of spheres using the CMM under test. Although the plate includes more spheres, we

eventually began to use only 12 spheres - a three by three grid in the plane and three on the spar - for our normal testing sequence. This represents our compromise between a more thorough sampling of the machine/workpiece geometry variation and a representative duration for a single execution of the measurement sequence. Measuring 12 spheres has a duration longer than would be found in transfer line applications (0.5-3 minutes), but shorter than many stand-alone inspection programs run on the shop floor (20-30 minutes.) The complete 12 sphere measurement sequence was repeated every 15 minutes.

Results Presentation

Finally, and perhaps most importantly, we report on our selection of methods for analyzing and presenting the large quantities of measurement data. We propose to use relatively simple quantities to characterize the thermal response of the workpiece-machine system and to estimate the thermal contribution to dimensional measurement uncertainty.

The record of the sphere centers is very analogous to a multi-point drift test, and the results of the test were initially presented in this manner. However, we concluded that the most meaningful (and easily explained) quantity is apparent variation in the distance between the spheres. This is analogous to a length measurement, whose variation with temperature is most easily understood. Also, it is analogous to measuring ball bars. Hence, reporting sphere center-to-center distance variations relates well to both widely used CMM performance standards. Further, by computing the distances between all pairs of spheres, of which there are 66 for 12 spheres, we mix together results from combinations of lengths (e.g., from crossed diagonals.) This provides sensitivity to changes in machine squareness, rotational errors, and straightnesses.

Having cleverly "reduced" the 36 raw coordinate variations to 66 time-varying distances, some further simplification is needed. Our approach is to scale all of the measured distances to a few standard lengths. In fact, for the data sets described here, all 66 distances were reduced to just two representative lengths. All distances between 100 mm and 300 mm, when initially measured at 20 C, were scaled up or down by the ratio of 200 mm to each initially measured distance. This effectively made all of these distances equal to 200 mm at 20 C. For lengths between 300 mm and 500 mm, the distances were scaled proportionally to 400 mm. The "standardized lengths" (200 mm and 400 mm) were then subtracted from all of the scaled time-varying distances. This produced two collections of thermal variations that were comparable within themselves. All of the distance variations for each standardized length could be plotted on a single graph, but just the centerline and envelope of the mass of data is much more informative. Consequently, we further reduced the results by determining the mean, minimum and maximum variations for each standard length at each measurement time. Typical results of this are presented as Figure 3.

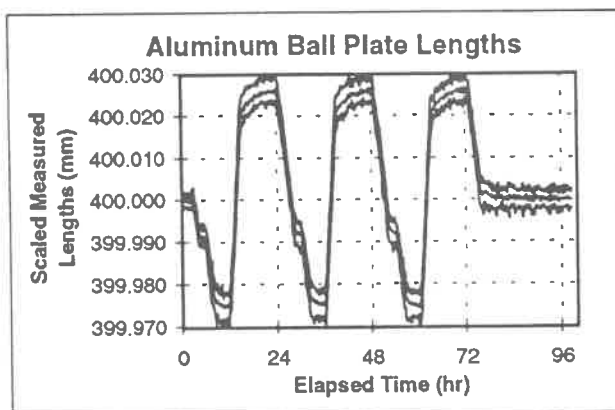


Figure 3. Standard Thermal Response Graph

ULTRAPRECISION GRINDING OF HIGH ASPECT RATIO MICROPARTS

Hans H. Gatzert¹, J. Christoph Maetzig², and Lars P. Ruhbach¹

¹ Institute for Microtechnology (imt), Hanover University, Hanover, Germany

² Now at Balzers Process-Systems, Alzenau, Germany

Abstract - Mechanical ultraprecision manufacturing technologies serve as an interesting alternative to fabricating structures by micromachining using thin film techniques. The main advantage of employing mechanical techniques is their capability of achieving a high aspect ratio between height and thickness of a microstructure. The paper describes investigations in how to accomplish the high aspect ratio profiling of an advanced ceramic workpiece. The test structure selected is a flexural resonator, the aspect ratio achieved is 50:1.

I. INTRODUCTION

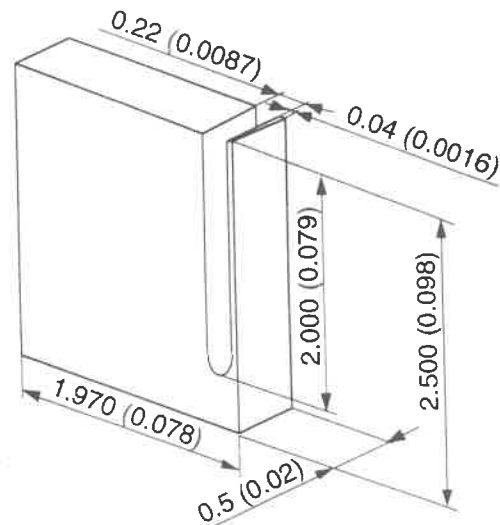
Mechanical ultraprecision machining of ceramic microparts has been spearheaded by magnetic read/write heads for rigid disk or tape drive applications [1, 2]. Mechanical ultraprecision manufacturing technologies provide for an interesting alternative to fabricating structures by micromachining using thin film techniques. While the latter in general is only capable of creating very thin structures, mechanical ultraprecision machining allows to build components with high aspect ratios, i.e. parts with a height substantially greater than their length and width.

To accomplish high aspect ratio ultraprecision grinding, three key conditions have to be met. First, the ceramic workpiece material has to be free of subsurface damages like microcracks or residual stresses induced during preceding machining operations. Furthermore, the profiling has to be by ductile mode grinding in order not to compromise the workpiece quality by subsurface damages induced during the dicing process. This machining mode is also a precondition for achieving an optimal sidewall surface and edge quality [3, 4]. Finally, the dicing blade needs to be stabilized properly to avoid wheel walk which would inflict damage to the structure. For investigations on how to achieve high aspect ratios of a mechanical structure with little subsurface damage, a flexural resonator has been selected as a test structure. The

mechanical processes applied are dicing and nanogrinding.

II. TEST SPECIMEN FOR HIGH ASPECT RATIO PROFILING

For this high aspect ratio microgrinding study, Altic, an advanced ceramic consisting of 70% Al_2O_3 and 30% TiC has been chosen as workpiece material. Altic has been used successfully as a wafer material for data recording thin film heads. Through its excellent mechanical properties, like a Vickers Hardness of 2,000 HV, a Bending Strength of 880 MPa, and a Young's Modulus of 390,000 MPa Altic is the ideal material for building the test specimen. The average grain size is $1.5 \mu\text{m}$ [5].



All dimensions in mm (in)

Fig. 1: Microresonator test structure.

Fig. 1 depicts the resonator test structure: a resonator tongue is created by cutting a groove into a ceramic block. Equation {1} describes the resonance frequency at the first bending mode of such a structure. The smaller the width of the tongue and the higher the aspect ratio between tongue length and tongue thickness chosen, the lower is the resonance frequency of such a system. To create a microresonator with a

resonance frequency at the upper audible range around and above 10 kHz, an aspect ratio of 50:1 is required. Therefore, such a microresonator is an excellent test object for high aspect ratio microprofiling.

$$f = \lambda^2 \cdot \frac{1}{2\pi} \sqrt{\frac{E \cdot h^2}{r \cdot 12 \cdot l^4}} \quad \{1\}$$

- l - Resonator Length (Height)
- r - Material Density
- h - Resonator Thickness (Width)
- f - Characteristic Frequency
- E - Young's Modulus

III. WORKPIECE SURFACE MACHINING

Cutting a block of advanced ceramic to the workpiece's outer dimensions typically involves surface grinding, which is mostly accomplished by brittle fracture mode cutting. This type of cutting causes a substantial amount of subsurface damages by inducing microcracks in the workpiece. In advanced ceramics, the depth of the subsurface layer having suffered damages during surface grinding typically is 20 μm [6]. A material having suffered this kind of subsurface damage will not lend itself to any consecutive ultraprecision machining unless the damaged layer is completely removed. The surface layer therefore has to be removed before ultraprecision profiling may be attempted. For removing the damaged layer, as well as for creating a high quality workpiece surface, each of the block's surfaces is machined by nanogrinding [7]. Nanogrinding is an ultraprecision grinding process developed by the imt and is based on lapping kinematics. The nanogrinding machines (Peter Wolters, Type 3R40) are installed under laminar flow to avoid any plate contamination. While lapping uses loose abrasive grain for material removal, for nanogrinding diamond grain is completely embedded into a soft metallic plate. The diamond summits are aligned in a plane, thus forming a diamond brush. The workpiece is machined by ductile regime cutting, the depth of cut of an individual grain is in a range of 1 to 5 nm. The surface roughnesses of nanoground AltiC blocks are $R_a = 0.75 \text{ nm}$ / $R_z = 8 \text{ nm}$. Fig. 2 depicts an AFM image of an AltiC block's nanoground surface. With the AltiC block's damaged surface layer removed by nanogrinding, the workpiece is now properly

prepared for a high aspect ratio profiling of the resonator tongue.

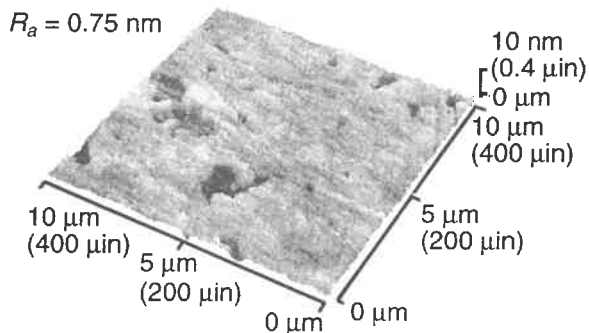


Fig. 2: Nanoground AltiC surface, surface topography measurement by AFM.

IV. HIGH ASPECT RATIO ULTRAPRECISION PROFILING

The grinding machine used in these tests is a single wheel dicing saw (Disco DAC-550) equipped with an air bearing spindle and an air slide, which allows minimum indexing step widths of 0.5 μm . To make full use of the machine accuracy, machine and tool deformation is minimized by holding both the room and the water coolant at a constant temperature of $21^\circ\text{C} \pm 1^\circ\text{C}$ ($70^\circ\text{F} \pm 1.8^\circ\text{F}$). The coolant is tap water with an additive to prevent rust and algae growth (Castrol Syntilo 9822). The cutting tools employed for the test specimen's fabrication are hubless type diamond blades. These blades may be resin bonded or metal bonded. In cutting hard and brittle ceramics, resin bonded wheels are superior to metal bonded ones. When the proper resin hardness is chosen, the diamond grains break out rather than get dull thus always providing new sharp diamond grains for cutting. Additionally, this process provides a space for the chips cut off from the work piece. Metal bonded blades, on the other hand, have a better radial wear behavior and therefore are mostly preferred for production. Since the diamond grains stay in the blade for a substantially greater period of time, they don't feature the superb self sharpening properties the resin bonded wheels excel in. For these investigations, blade accuracy is considered more important than blade durability; therefore resin bonded wheels have been chosen. As discussed previously, brittle fracture cutting is the most common cutting mode when surface grinding advanced ceramics. This also holds true for dicing and

profiling. Due to the subsurface damages this cutting mode is inducing, it is not feasible to use this particular cutting mode for ultraprecision machining. In case of a flexural resonator, subsurface damages also are not permitted since they will cause tongue breakage during operation. However, advanced ceramics may also be cut in a ductile mode. Reaching the ductile mode grinding regime requires a substantially higher induced energy than when cutting in the brittle fracture grinding mode. Furthermore, cutting conditions have to be chosen in a way that a submicron depth of cut is achieved. Blades used for microprofiling have to be very thin, which compromises blade stiffness. The blade thickness chosen for our tests is 200 μm (0.008 in). Insufficient blade thickness results in wheel walk: the groove the blade is cutting is not straight but veers off to one side. To minimize detrimental effects due to an insufficient blade thickness, two steps may be taken. The first measure is to minimize the blade exposure. This is accomplished by choosing a flange diameter minutely smaller than the difference between the blade diameter and the desired depth of cut. In our case, to achieve a depth of cut of 2 mm (0.080 in), a flange diameter of 50.6 mm (1.99 in) and a blade outer diameter of 55 mm (2.17 in) is used. The second one is to properly guide the blade. In our case, to provide for an optimal blade guidance when cutting the groove to profile the delicate resonator tongue, dummy blocks are attached to the workpiece at both the blade entry and blade exit side as shown in Fig. 3.

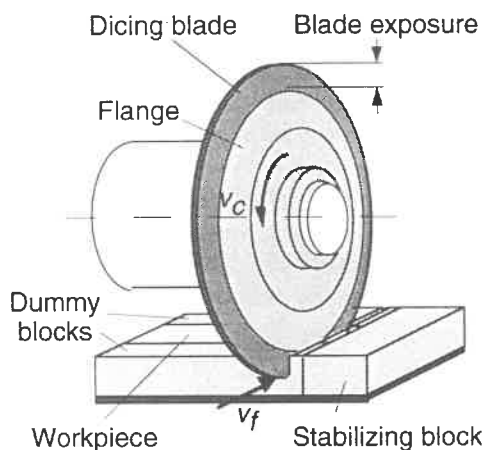


Fig. 3: Microprofiling machining set-up.

Furthermore, a stabilizer block stiffens up the tongue side of the workpiece to avoid is

breakage during the cutting process. An adhesive (Aremco Crystal Bond 509) is used to attach workpiece and blocks to each other and to a carrier. For the majority of the tests, a resin bonded blade (Microkerf: 2.187-8-15H) with hard resin binder, a diamond grain size of 15 μm , and diamond concentration C100 (4.39 carat/ cm^3) has been chosen. This choice is based on previous investigations on how to achieve optimal cutting conditions when machining Altic [1]. In most cases, cutting or profiling of ceramics is done in a single run. However, cuts with a cutting depth greater than 1 mm (0.04 in) caused excessive forces at the workpiece resulting in a tongue breakage. Therefore, experiments were made to cut the groove in two steps, each step cutting half the groove depth. With this approach excellent results are achieved.

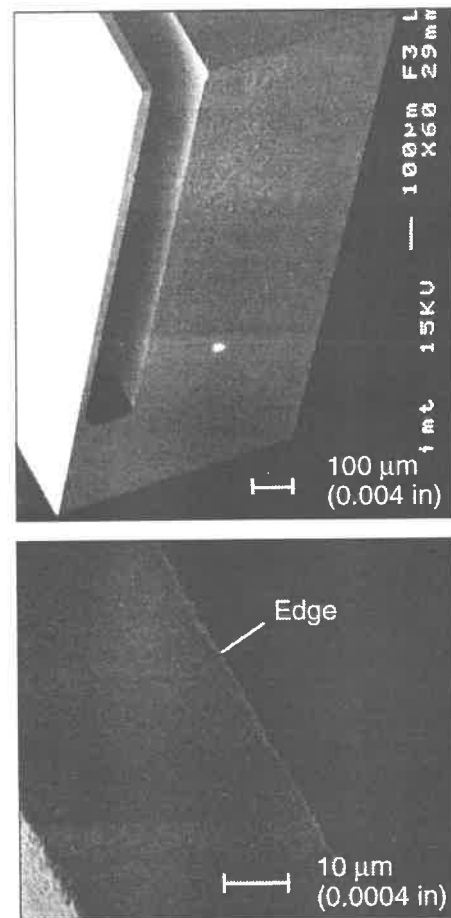


Fig. 4: SEM micrograph of resonator after tongue profiling/SEM micrograph of resonator tongue edge.

Fig. 4 depicts on its left side a micrograph of the resonator after cutting the tongue and on its right side the kerf's edge quality. The aspect ratio between tongue height and tongue thickness is 50:1. The cutting velocity is 86.4 m/s (30,000 RPM), the feed rate $v_f = 3$ mm/min (0.12 in/min). Before its use, a dicing blade has to be trued and sharpened. In general, a square shape of the dicing blade is desirable. However, for the given test structure a rounded kerf ground is required to avoid a stress concentration at the tongue's bottom due to a notch effect, which is likely to lead to breakage of the microstructure while oscillating in its sensor application. Therefore, the blade edges are rounded in a truing process using a 45 μm grain size SiC block (Disco BGCA0113) moved circumferentially around the blade. For dressing a 35 μm grain size SiC block (Lunzer M50) is used. Fig. 5 presents the kerf's bottom created by such a blade.

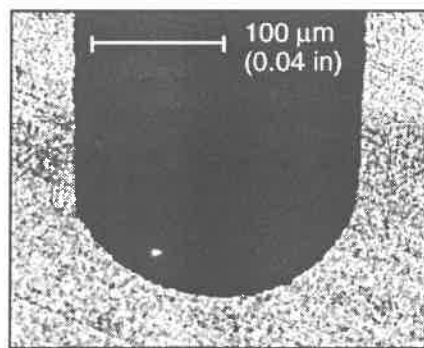


Fig. 5: Micrograph of the kerf radius.

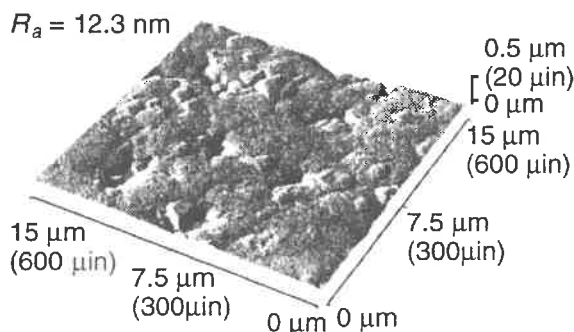


Fig. 6: Surface topography measurement by AFM of the kerf sidewall.

Fig. 6 depicts an AFM surface topography plot of the kerf's sidewall. Due to its contact with

the blade's side wall, the kerf's sidewall is prone to suffering damages. There is also an indication that plastic deformation due to side wall friction causes a smoothening by plastic deformation.

V. CONCLUSION

A combination of nanogrinding and ultra-precision profiling allows the creation of a microstructure with high aspect ratios between structure height and structure thickness. Nanogrinding is used to remove subsurface damages at the workpieces' outer surfaces. The profiling result depends on choosing the machining parameters to achieve the ductile regime cutting, as well as avoiding wheel walk by properly stabilizing blade and workpiece. As a test structure, a flexural resonator has been machined, an aspect ratio between resonator tongue height and thickness of 50:1 has been accomplished by a two step cut with a depth of cut of 1 mm (0.040 in) each.

VI. REFERENCES

- [1] Gatzen, H. H.; Maetzig, J. C.; Schwabe, M. K.: "Precision Machining of Rigid Disk Drive Heads Sliders", IEEE Transactions of Magnetics, (1996), 32(3), p. 1843-1849
- [2] Gatzen, H. H.: "Rigid Disk Recording Head Precision Engineering Requirements", 9th IPES/ 4th UME Braunschweig, Germany, (1997), Vol.1, p. 10-16
- [3] Bifano, T. G.: "Ductile Regime Grinding: A New Technology for Machining Brittle Materials", Transactions of the ASME, (1991), Vol. 112, p. 184-189
- [4] Komanduri, R.: "On Material Removal Mechanisms in Finishing of Advanced Ceramics and Glasses", Annals of the CIRP, (1996), Vol. 45/1, p. 509-513
- [5] Source: Sumitomo Special Metals Co. Ltd. Japan
- [6] Uhlmann, E. G.: "Tiefschleifen hochfester keramischer Werkstoffe" ("Creep Feed Grinding of Advanced Ceramics"), PhD Thesis, Berlin University (TU), (1994), p. 216-219
- [7] Gatzen, H. H.; Maetzig, J. C.: "Nanogrinding", Received for Publication in: Journal of the American Society for Precision Engineering, (1997)

Direct Fabrication of Deep X-ray Lithography Masks by Micromechanical Milling

Craig Friedrich ^a, Philip Coane ^b, Jost Goettert ^b, Niranjan Gopinathin ^c

^a Michigan Technological University (MEEM), Houghton, MI

^b Louisiana Tech University (IfM), Ruston, LA

^c Read-Rite, Inc., Milpitas, CA

Introduction

Micromechanical milling has been shown to be a rapid and versatile method for the fabrication of microstructures in polymethyl methacrylate (PMMA) with lateral dimensions down to 4 micrometers [1-3]. These structures were micromilled in PMMA using custom-fabricated micromilling tools with a diameter of 22 micrometers [4]. To make the micromilling process more cost effective, it has been adapted for the fabrication of deep x-ray lithography masks. In deep x-ray lithography, the absorber pattern is transferred into the PMMA photoresist using x-rays as the illumination source.

Conventional deep x-ray lithography (DXRL) mask fabrication

Deep x-ray lithography masks must provide high contrast between the transmitted x-rays, through the *membrane*, and the absorbed x-rays, in the *absorber*, because exposures can take many hours. Membrane materials must offer high transparency to x-rays and are made of low absorption materials (low absolute absorption and thickness) such as beryllium, carbon, silicon, titanium, silicon nitride, or boron nitride. The higher atomic number membrane materials are usually a chemically etched, thin membrane, where the low atomic number materials are normally thicker. The absorber pattern is made of materials with high x-ray absorption such as tungsten, gold, or tantalum which have a high absorption coefficient for x-rays in the range of 2 keV to 30 keV.

Micromechanically Milled Masks

Micromilled masks have the same general characteristics as the lithography counterparts, but the materials and dimensions must be altered to allow for mechanical machining. The depth of the features depends on the diameter of the tool used and the desired result. Because the absorber thickness in most x-ray masks is on the order of 10 micrometers, the micromilling process may be a direct and rapid method for x-ray mask fabrication.

Micromilled mask materials

A mask structure capable of being machined must have a relatively thick *substrate* (transmission "membrane") because the flatness of the substrate, coupled with variations in the thickness, preclude the use of thin membranes as are commonly used in conventional masks. At the same time, the absorber deposited onto the substrate must be sufficiently opaque to x-rays, must be machinable with minimal cutting tool wear and adequate dimensional control, and must be deposited onto the substrate with adequate adhesion so as not to delaminate in the presence of the machining forces.

The first work used a single material as both the substrate and absorber. Aluminum was chosen and preliminary tests indicated a thickness of 6 micrometers for the substrate, and a total of 40 micrometers for the absorber gave reasonable contrast when exposed on the Institute for Micromanufacturing (IfM) beamline at the Center for Advanced Microstructures and Devices (CAMD). The drawback which precluded further development of this architecture was the time and care required to mill the aluminum foil down to a minimum thickness of 6 micrometers. Although attainable, it took great care and this detracted from the rapid intent of the process.

The next step was to investigate a low absorption substrate and a high absorption, sputtered

absorber. Because the substrate was subjected to machining, beryllium was not considered because of the toxicity of the oxide. The lowest atomic number structural material available was carbon. Test exposures were made with varying thicknesses of commercially available (Goodfellow) graphite sheets. It was found that a thickness of less than 250 micrometers gave sufficient transmission of the x-rays to allow for timely exposures (8 hours or less).

Work on an appropriate absorber was more extensive. Gold was the first material used, however sputtered gold was too ductile to leave a cleanly machined pattern. Other sputtered single layer absorbers were then tried. These were tantalum, titanium, tungsten, and molybdenum and were sputtered to a thickness of 5 micrometers (8 micrometers for molybdenum) over a 1000 - 2500 Angstrom- thick layer of chromium. The absorber adhesion was good, but cutting tool wear was excessive, particularly with the tantalum and tungsten. In addition, a thickness of 5 micrometers did not provide sufficient absorption and low contrast was observed in the exposed structures.

Because of the rapid tool wear, it was decided to use a thick layer of gold as the primary absorber and a thin layer of a harder metal on top to help reduce the burring of the gold. The best combination was gold covered with titanium. The sputtered gold layer was 8 micrometers thick and the sputtered titanium layer was 3 micrometers. A series of patterns, covering a total of more than 625 square millimeters, were micromilled to determine the quality of this mask combination. The total time required to mill the extensive patterns was about 4 hours. The machining parameters were a single axial pass at a depth of 36 micrometers (11 micrometer-thick absorber and 25 micrometers of overcut to account for any lack of blank flatness), a spindle speed of 10,000 rpm, and a feedrate of 30.5 micrometers per second with a 100-micrometer milling tool of tungsten carbide. The mask was exposed at CAMD and two results are shown in **Figures 1 and 2**.

Mask Absorber Quantification

The second phase of the study was to determine the precision of mask absorber features and the exposed pattern in the PMMA photoresist. There are two primary biases in the mask/exposure process. The first is the precision of the mask absorber definition and the second is the bias in transferring the absorber pattern into the photoresist.

The micromilling tools are held in a precision collet in the milling machine spindle. The tool has radial runout in the collet and there is also eccentricity between the tool shank and the cutting edges. To accurately machine trenches, which will leave narrow absorber features, this runout must be eliminated to the extent possible and the remaining error must be quantified so it can be compensated. Much of the runout can be removed by "tapping in" the tool using a video microscope. Typically, the remaining runout after this procedure is 3 to 5 micrometers, based on experience.

To quantify and compensate the remaining runout required two test cuts to be made in the mask blank, well away from the area to be patterned. Two parallel trenches, each 3 mm long, were milled into the absorber. The remaining absorber width was then measured with a WYKO RST-Plus at 80X. The wall width was measured at 20 points along the length and the mean was used to determine the effective tool diameter which was 109 micrometers (for a nominal tool diameter of 100 micrometers), with a standard deviation of 0.9 micrometers.

The effective tool diameter was then used as a process bias and five similar trenches were machined with the same feed, speed, and axial depth with a varying center spacing. The desired absorber widths between the trenches were 50, 20, 15, and 10 micrometers. Each absorber wall was measured at 20 points along the length and the results are shown in **Table 1**.

Exposure Quantification

The micromilled mask was then used to expose 300 micrometer-thick PMMA at the CAMD synchrotron under the same conditions mentioned previously. The result of the exposure and development is shown in **Figure 3**. The exposed wall widths were measured as previously described and the results are shown in **Table 2**. It can be seen that the exposed widths are consistently narrower than the measured

absorber widths. The mask absorber wall widths were again checked and it was found that the milling procedure left a slight taper in the absorber, particularly the gold, whereby the absorber thickness was less at the edge of the wall and quickly grew to full thickness a micrometer, or so, away from the wall edge. This variable thickness went unnoticed in the preliminary absorber width measurements but was observed in a more detailed inspection of the scanning electron micrographs. This created a second process bias which must be quantified to determine how repeatable it is and if the amount of taper is a function of the machining parameters (feed and speed).

One of the exposed samples was sectioned to determine the roughness in the vertical walls. A series of two-dimensional line scans were taken which covered the entire 3 mm length of the walls. The average mean roughness (Ra) was 0.17 micrometers with a standard deviation of 0.03, and the average peak-to-valley (Rt) was 1.42 micrometers with a standard deviation of 0.24. The relatively large peak-to-valley roughness was due to residual burring in the absorber which was transferred into the photoresist, but the relatively low average roughness indicated the burrs were random and occasional, and not overly prevalent. Similar wall roughness was measured in exposed PMMA using a conventional gold mask and the average mean roughness (Ra) was 0.08 micrometers with a standard deviation of 0.015, and the peak-to-valley roughness (Rt) was 0.51 micrometers with a standard deviation of 0.1 micrometers [6].

Acknowledgments

This work was supported by the Institute for Micromanufacturing under NSF grants OSR-9550481 and EEC-9420600 and the State of Louisiana under grant NSF/LEQSF(1995-98)-SI-01. Special thanks to Joe Blakesley and Jubin Shah for sputtering the absorbers.

References

1. Friedrich, C. and Vasile, M., "The micromilling process for high aspect ratio microstructures", *Microsystem Technologies* (2)3, pp.144-148, 1996.
2. Vasile, M., Friedrich, C., Kikkeri, B., and McElhannon, R., "Micron-scale machining: tool fabrication and initial results", *Precision Engineering* (19)2/3, pp.180-186, 1996.
3. Friedrich, C., Coane, P., Goettert, J., and Gopinathin, N., "Metrology and quantification of micro-milled x-ray masks and exposures, *SPIE Vol. 3048* (in press), 1997.
4. Friedrich, C. and Vasile, M., "Development of the micromilling process for high aspect ratio microstructures", *J. MEMS* (5)1, pp.33-38, 1996.
5. Coane, P. and Friedrich, C., "Fabrication of composite x-ray masks by micromilling", *SPIE Vol. 2880*, pp.130-141, 1996.
6. Gopinathin, N., *Direct fabrication of x-ray masks for liga by micromilling*, MS Thesis, Louisiana Tech University, Ruston, LA, 1997.

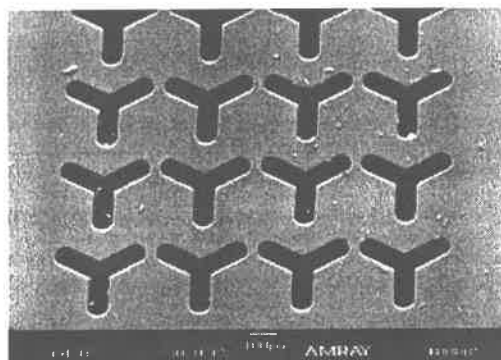


Figure 1. Portion of exposure in PMMA through Au-Ti micromilled mask. Scale bar = 100 μ m.

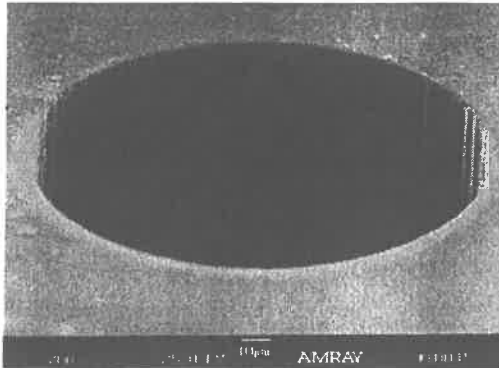


Figure 2. Portion of exposure in PMMA through Au-Ti micromilled mask. Scale bar = 10 μm .

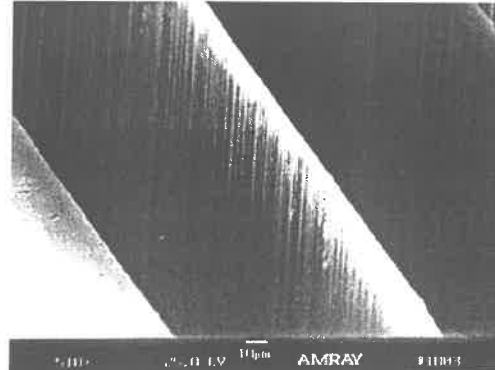


Figure 3. Closeup of 10-micrometer wide by 300-micrometer high exposed and developed wall. Scale bar = 10 μm .

Table 1. Mask absorber width accuracy using tool runout characterization.

Nominal absorber wall width (μm)	Measured absorber wall width (μm)	Standard deviation (μm)
50	50.4	0.5
20	20.7	0.4
15	15.4	0.8
10	11.1	0.5

Table 2. Exposed wall width accuracy in PMMA.
Large standard deviation is due to roughness from mask burrs.

Nominal absorber wall width (μm)	Measured PMMA wall width (μm)	Standard deviation (μm)
50	48.2	1.4
20	16.0	1.8
15	13.6	1.3
10	9.6	0.8

THE MECHANISM OF MATERIAL REMOVAL IN GRINDING AND POLISHING SILICON MONO-CRYSTALS ON THE NANOMETRE SCALE

L. C. Zhang, H. Tanaka and I. Zarudi

Department of Mechanical and Mechatronic Engineering
The University of Sydney, NSW 2006, Australia

1. Introduction

Grinding and polishing are important machining processes in preparing high quality components of silicon mono-crystals for integrated circuits and electronic devices. These machining processes must be applied carefully to avoid surface and subsurface damages. It has been widely accepted that a 'damage-free' grinding or polishing of ceramics, such as alumina, can be achieved with certain processing conditions, under which sufficient number of independent slip systems are activated [1-3]. However, the mechanism of material removal associated with damage-free grinding or polishing of silicon has not been understood well. This paper aims to explore the mechanism using molecular dynamics (MD) simulation and experimental analysis.

2. Method

2.1 Molecular dynamics modelling

To reveal dynamically the deformation of silicon mono-crystals subjected to grinding and polishing on the nanometre scale, the three-dimensional molecular dynamics method is used to understand the deformation of the atomic lattice of silicon.

In grinding, an abrasive cutting edge is fixed on a grinding wheel so that the cutting of a diamond tip can be simulated approximately by a cutting process, as shown in Fig. 1, with a hemispherical diamond abrasive of radius R moving from the right to the left with the peripheral speed of a grinding wheel V_c . In polishing, however, abrasives are in a slurry and thus the motion of each abrasive has both a translation of speed V_t and a self-rotation of peripheral speed V_r during polishing. Such abrasive motion is simulated by the molecular dynamics model shown in Fig. 2 with an abrasive ball of diamond of radius R moving horizontally (translation) and rotating about its centre independently at the same time.

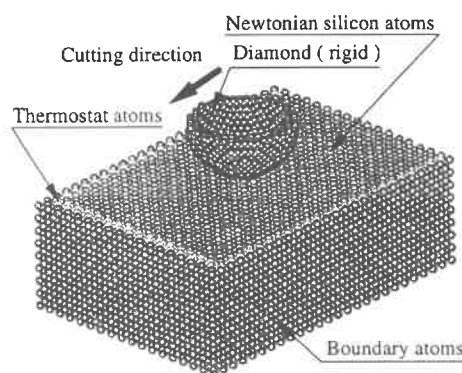


Figure 1 The initial MD model of cutting the (100) surface of a silicon mono-crystal. The indentation depth of the diamond tip is d_i .

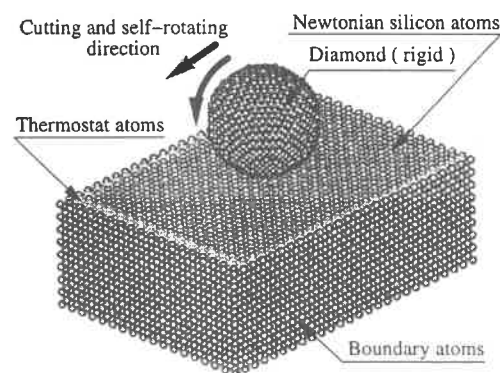


Figure 2 The initial MD model of polishing the (100) surface of a silicon mono-crystal. The indentation depth of the diamond ball is d_i .

To avoid the boundary effect, the dimension of the moving control volume (MCV) of the silicon specimen was taken as $9.23 \text{ nm} \times 13.57 \text{ nm} \times 4.34 \text{ nm}$. As usual [4-5], to eliminate the rigid body motion of the specimen and guarantee a reasonable heat conduction outwards the MCV,

layers of boundary atoms, which are fixed to the space, and layers of thermostat atoms, which absorb the heat at the MCV boundaries, are arranged to surround the Newtonian atoms of silicon in the MCV except its top surface (100) that is subjected to abrasion. According to the atomic structure of silicon, Tersoff potential was used to describe the interactions among silicon atoms but Morse potential was employed to reflect those among silicon and diamond atoms. Detailed information about Tersoff and Morse potential functions can be found in Refs. [3-6].

Another important factor in a successful molecular dynamics modelling is the reliability of the relationship between the kinetic energy and temperature of an atom. Currently, there are three models available [7], *ie*, Dulong-Petit model, which takes into account the independent lattice vibration, Einstein model, which is based on the consideration of the single characteristic frequency, and Debye model, which involves a range of frequencies. Our comparison with experiments shows that in the temperature regime encountered in the grinding and polishing processes to be discussed in this paper, Debye model is the best for silicon and Einstein model is the most applicable to diamond.

2.2 Experimental analysis

Experimental investigations were also carried out in two ways, *ie*, in polishing and grinding, to examine the variations of surface and subsurface microstructures of silicon mono-crystals with processing conditions and to understand the related mechanism revealed by MD analysis. In grinding, the (100) surfaces of silicon mono-crystals were ground in the $\langle 110 \rangle$ direction on a modified ultra-precision surface grinder, Minini Junior 90CNCM286 (loop stiffness = $80 \text{ N}/\mu\text{m}$), using a diamond wheel, SD4000L75BPF, of 305 mm in diameter rotating at 1500 rpm . The depth of cut of the grinding wheel was fixed at 100 nm but the table speed was varied from 0.02 m/min to 1 m/min to alter the nominal chip thickness. The coolant used was Syntilo 3 (99% water and 1 % mineral oil).

To prepare surfaces by polishing, the silicon mono-crystal specimens were first ground by loose abrasives of silicon carbide with a mean diameter of $50 \mu\text{m}$ and followed by a similar grinding with abrasives of $25 \mu\text{m}$ in diameter. After that, a series of fine polishing were carried out using alumina abrasives of mean diameters of $15 \mu\text{m}$, $9 \mu\text{m}$, $5 \mu\text{m}$ and $1 \mu\text{m}$, respectively. A finishing polishing was carried out by ultra-fine particles of 25 nm in diameter. The polishing pressure applied was about 7.5 kPa . The arm (400 mm long) and disk of the polisher rotated at 40 rpm and 60 rpm , respectively. The finishing polishing took about 10 hours.

The topography of ground and polished surfaces were examined by a high resolution scanning electron microscope, JSM-6000F. The subsurface structure was investigated by a Transmission Electron Microscope, EM430, and a Scanning Transmission Electron Microscope, VGHB601.

3. Results

It found that there always existed a thin layer of amorphous silicon in a ground subsurface, as shown in Fig. 3. The thickness of the layer decreased with decreasing the nominal chip thickness. Dislocations of very low density were activated in the crystal silicon from the interface between the amorphous layer and the crystal substrate (Fig. 3a). When the nominal chip thickness or the depth of cut of cutting edges became extremely small, in the case of ultra-fine polishing in particular, no dislocation appeared. The residual amorphous layer became extremely thin and could not be detected even under a very high magnification shown in Fig. 4. In this case, 'truly' damage-free workpieces with perfect crystal structure were obtained. The mechanism of this phenomenon can be clearly elucidated by the atomic lattice deformation presented by the

molecular dynamics analysis. As shown in Figs. 5 and 6, the deformation of a silicon mono-crystal under grinding or polishing has the following distinct features:

(i) Phase transformation, from the original crystal structure of silicon to its amorphous state, happens almost simultaneously with the cutting edge penetration.

(ii) Chipping always takes place within the amorphous layer regardless of the appearance of dislocations in the crystal region. This means that the mechanism of material removal in grinding and polishing silicon mono-crystals is due to the viscous flow in the amorphous zone.

(iii) Re-crystallisation always occurs in the amorphous zone behind the cutting edge, from the bottom of the amorphous layer to the surface. As a result, the residual amorphous layer is always thinner. Under certain processing conditions (*eg*, small depth of cut and certain combination of translating and self-rotating speeds, *etc*), only a few atomic layer undergoes amorphous change during abrasion and a reverse phase transformation from amorphous to crystal may take place completely after abrasion. In this case, there will be no residual amorphous layer in the machined subsurface. In other words, a perfect crystal subsurface without any dislocations and amorphous change is generated.

(iv) In polishing, surface roughness may be improved by abrasive-rolling induced phase transformation without material removal. With an abrasive-indentation depth on the atomic scale, material removal may only occur *via* adhesion when abrasive-workmaterial interactions become sufficiently strong.

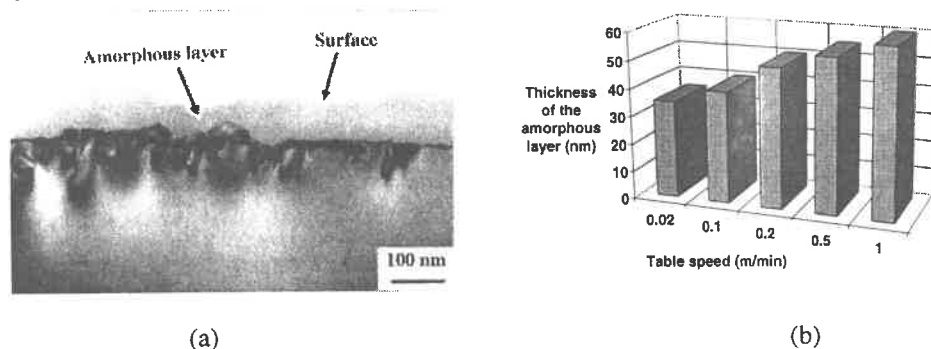


Figure 3 Subsurface structure of a silicon mono-crystal after grinding, (a) phase change and dislocations (table speed = 0.02 m/min), (b) Thickness of amorphous layer vs table speed.

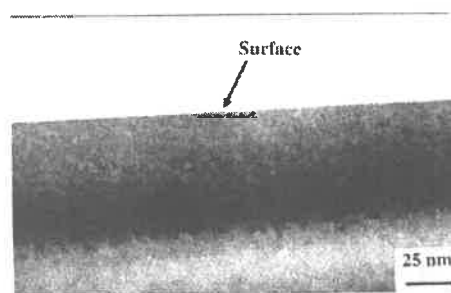


Figure 4 The 'perfect' subsurface structure of crystal obtained by the ultra-fine polishing process. The polished surface is (100). No dislocation exists. The residual amorphous layer cannot be detected. The ultra-fine polishing conditions used are presented in section 2.2 of this paper.

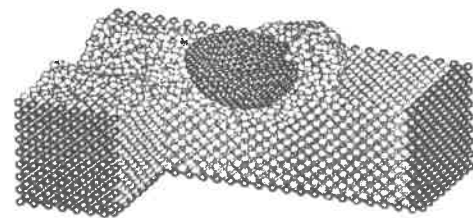


Figure 5 Chipping via viscous flow in the amorphous layer when cutting the (100) surface of a silicon mono-crystal. Note the thickness variation of the amorphous layer underneath and behind the diamond tip. Note also that no dislocations are generated. Cutting conditions are: $V_c = 200$ m/s, $d_i = 0.99$ nm and $R = 2.1$ nm.

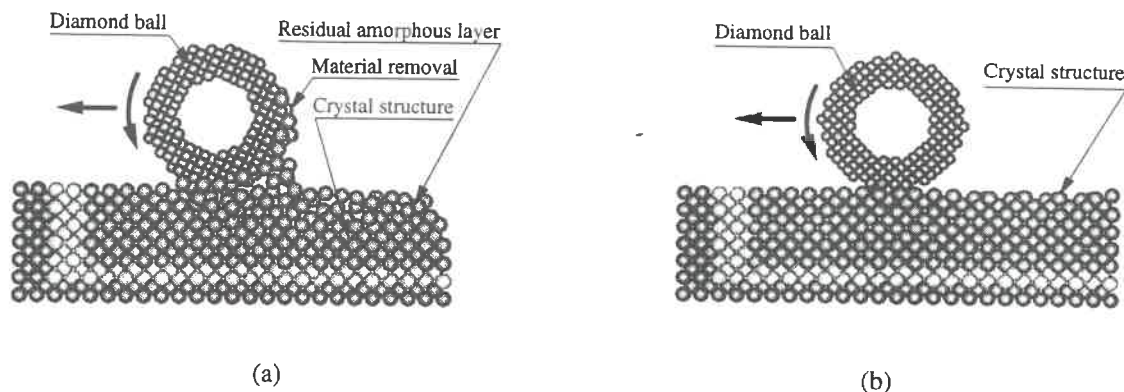


Figure 6 A cross-sectional view of the instant change of subsurface structure in silicon during polishing, (a) with a residual amorphous layer (polishing conditions: $R = 1.07 \text{ nm}$, $d_i = 0.09 \text{ nm}$, $V_t = 100 \text{ m/s}$ and $V_r = 50 \text{ m/s}$), (b) without any residual amorphous deformation or dislocations ($R = 1.07 \text{ nm}$, $d_i = 0.09 \text{ nm}$, $V_t = 100 \text{ m/s}$ and $V_r = 100 \text{ m/s}$).

4. Conclusions

According to the above analysis, we obtain the following conclusions:

(1) The material removal of grinding and polishing silicon mono-crystals on the nanometre scale is due to the viscous flow within the amorphous layer. Below a certain limit of abrasive indentation depth (on the atomic scale), grinding cannot remove material, but polishing can do through the mechanism of adhesion when abrasive-workmaterial interaction is sufficiently strong.

(2) Dislocations are not important to material removal. If the depth of cut of each cutting edge is sufficiently small, a 'perfect' crystal subsurface can be produced.

(3) In nano-grinding and nano-polishing silicon mono-crystals, deformation via amorphous phase transformation and its associated viscous flow is more favourable than that via dislocation.

Acknowledgment

The authors would like to thank the continuous financial support from the Australian Research Council to this research. Dr Nicole Bordes at the Vislab of Sydney University helped in MD visualisation and Mr. A Leistner at CSIRO Lindfield helped in polishing specimens.

References

- [1] I. Zarudi, L. C. Zhang and Y-W Mai, *J Materials Science*, **31** (1996) 905-914.
- [2] I. Zarudi and L C Zhang, in: *Procs 10th ASPE Annual Meeting*, Austin, Texas, USA (1995) pp.163-166.
- [3] I. Zarudi, L. C. Zhang and D. Cockayne, submitted to: *J Materials Processing Tech* (1997).
- [4] L C Zhang and H Tanaka, Towards a deeper understanding of friction and wear on the atomic scale: a molecular dynamics analysis, *Wear*, in press (1997).
- [5] H Tanaka and L C Zhang, in: *Progress of Cutting and Grinding*, edited by N Narutaki, et al, JSPE, Osaka, Japan (1997) pp.262-266.
- [6] H. Tanaka and L C Zhang, in: *Advances in Abrasive Technology*, edited by L C Zhang and N Yasunaga, World Scientific, Singapore (1997) pp.43-47.
- [7] C Kittel, *Introduction to Solid State Physics*, 7th edn, John Wiley & Sons Inc, New York (1996).

A Basic Study on Nanometer Cutting of Brittle Materials

Kohsaku ASAMOTO*, Katsushi FURUTANI, Naotake MOHRI

Toyota Technological Institute

12-1, Hisakata 2-chome, Tempaku-ku, Nagoya 468 JAPAN

Phone: +81-52-802-1111, Fax: +81-52-809-1721

* *Koyo Seiko Co. Ltd.* at present

Key words: nanometer cutting, brittle material, atomic force microscope (AFM), cantilever, frictional force microscope (FFM), wear, single crystal silicon

1. Introduction

In recent years, micromachining technology has been utilized in various fields [1]. The progress of technologies will require more precise machining.

Fine cutting with an atomic force microscope (AFM) suitable for observing in minute area has been studied by experiments and simulations [2][3]. In these studies, mechanical removal process is performed with the probe of the AFM for a cutting tool. Nanometer order shapes have been obtained by cutting in air. Single crystal silicon is usually used for a workpiece in these experiments because of its homogeneity. By cutting with the depth less than sub-micrometer order, silicon can be machined in the ductile mode with continuous chips [4].

The authors define the cutting brittle materials in the ductile mode with the nanometer roughness as the nanometer cutting.

In this paper, a measurement method of the wear of the probe tip is proposed at first. Then the influence of cutting conditions on the machinability of the nanometer cutting and the tendency of tool wear are experimentally investigated.

2. Determination of machining and measurement load

In the experiments, an AFM (SPM 9500 made by Shimadzu) is used to machine a workpiece and observe its shapes. A probe tip is made from Si_3N_4 . It has a pyramid shape with the nominal radius of 20 nm. Since a cantilever is made by a leaf spring, the load of the tip on a workpiece can be changed by elevating the workpiece with a piezo scanner.

Before machining shapes, a force curve is measured to determine the machining and the measurement loads.

Fig. 1 shows an example of the force curve measured in air. The vertical and horizontal axes indicate the output of a divided photodiode and the applied voltage to the piezo scanner used for scanning motion in the xy plane, respectively. The output of the photodiode is proportional to the deflection of the cantilever in the z direction which depends on the load act on the tip. The positive deflection corresponds to the repulsive force between the tip and the specimen. The tip approaches the specimen when the applied voltage to the piezo is decreased. As the distance between the tip and the specimen is changed, the load act on the cantilever is changed by the capillary phenomena on H_2O molecule layer, the atomic force and the meniscus force as shown in Fig. 1. Since the displacement of the specimen is equal to the deflection of the cantilever by the atomic force on the linear part of the force curve, the load act on the cantilever can be

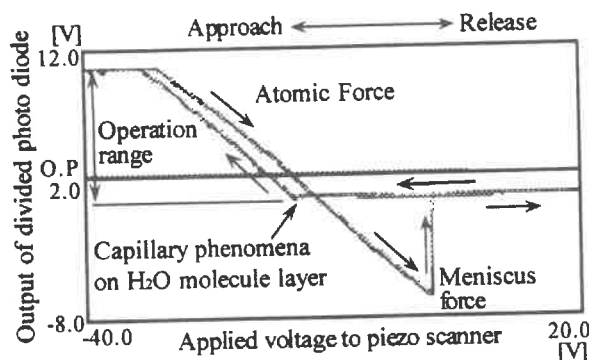


Fig. 1 Example of force curve in air

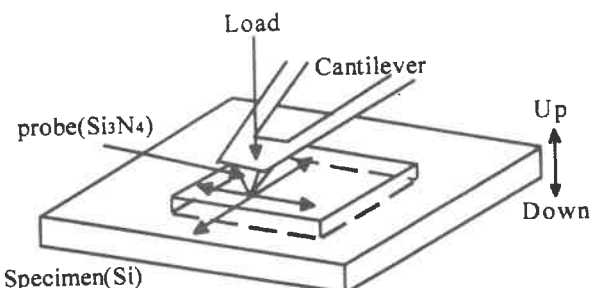


Fig. 2 Method of scratch machining with tip of AFM

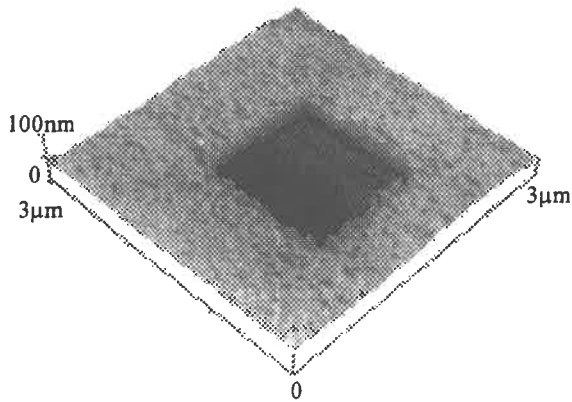


Fig. 3 Experimental result of machining measured with AFM

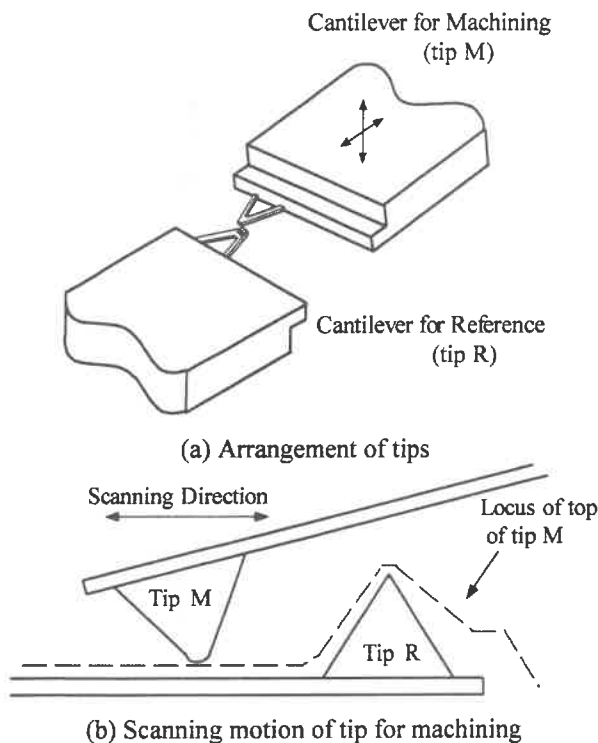


Fig. 4 Tip shape measurement by scanning over another tip

calculated by using the given spring constant [5]. By using this relation, the machining and measurement loads are determined as 196 nN and 9.8 nN, respectively.

Fig. 2 illustrates a method of scratch machining with the tip of the AFM. By scanning the tip in the x and y directions, shapes can be machined in the z direction.

Fig. 3 shows an experimental result of machining measured with the AFM. A silicon wafer cut along (100) plane is used for a workpiece after the RCA

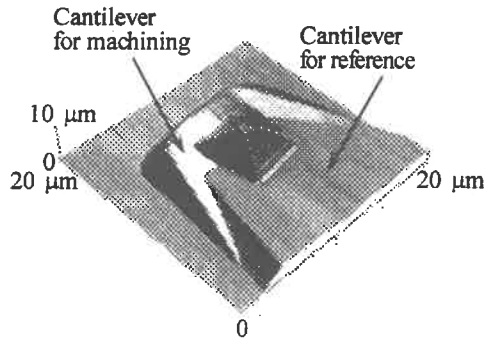


Fig. 5 Example of measurement of tip

cleaning. The surface with nanometer roughness is obtained by scratching.

3. Measurement of tip wear

3.1 Measurement method

The sharpness of a cutting tool affects the roughness of the machined surface in macro mechanical process.

Since a tip of the AFM is used for the cutting tool in this study, the tip shape should be measured after each scratching to avoid the influence of the tool shape.

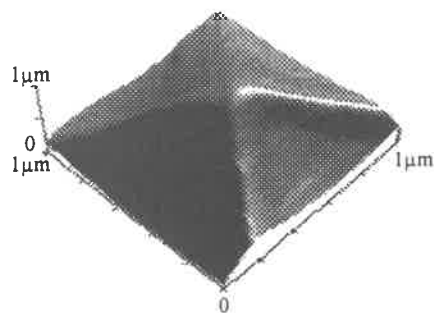
In general, a specimen with uniform protrusions is used for a reference to estimate the wear of the tip. In this estimation, the decrease of the spacial frequency indicates the wear of the tip. However, the shape of the tip can not be directly measured by this method.

Fig. 4 illustrates the proposed measurement method of the tip shape. Two tips are faced each other: the tip for machining (tip M) and the tip for the reference (tip R). The sum of their curvature radii can be obtained by scanning the tip M over the tip R. When the radius of the tip R is much smaller than the that of the tip M, the radius of the tip M can be measured approximately. The progress of the tip wear can be quantitatively measured by observing the shape of the tips and comparing the shape with the initial shape.

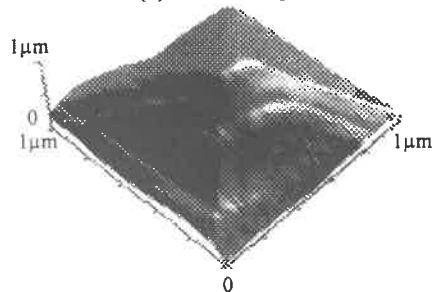
Fig. 5 shows an example of the measurement of the tip. The odd shape of the tip M is observed near its boundary by the interaction between the tips M and R. However, there is no problem on the measurement of the top of the tip. The progress of the tip wear during machining can be observed as shown in Fig. 6.

3.2 Results of measurement

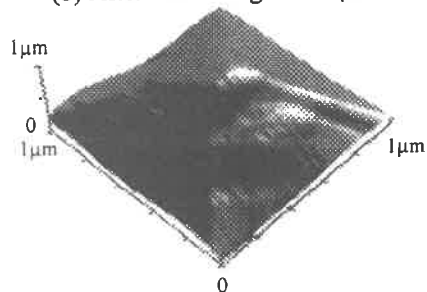
The measurement of the amount of the tip wear by the proposed method is introduced. Fig. 7 illustrates the principle of the tip wear measurement. At first, the



(a) Initial shape



(b) After machining $5 \times 10^4 \mu\text{m}^2$



(b) After machining $1 \times 10^5 \mu\text{m}^2$

Fig. 6 Progress of tip wear during machining

area of a horizontal of the cross section A on a horizontal plane and the volume of the pyramid above the plane are calculated as the reference area and initial volume of the tip, respectively. After n th machining, the n th volume of the pyramid is calculated with the bottom area B which has the same area as the section A. Then the tip wear can be evaluated by comparing n th volume with the initial volume. Since the origin of measured shapes with AFMs always drifts and this method does not depend on the relative coordinates, the drift of the origin does not affect the measured amount of the wear.

Fig. 8 shows an example of the progress of the tip wear. It can be observed by the proposed method that the tip wear increases with a constant rate during machining. The adhesion of the chip on the tip causes the distribution of the measured wear. The distribution can be reduced by removing the chip.

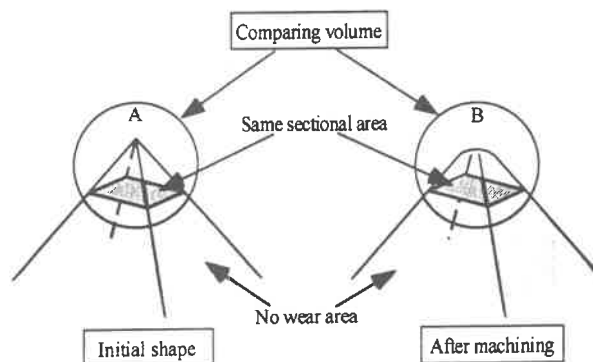


Fig. 7 Principle of tip wear measurement

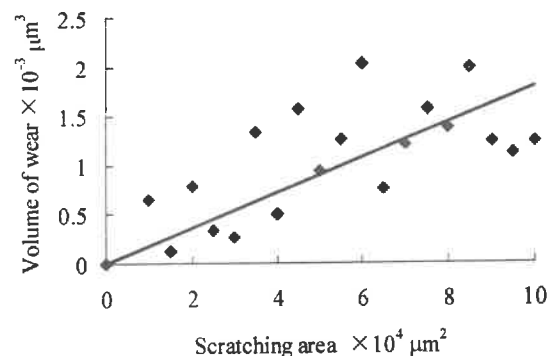


Fig. 8 Progress of tip wear

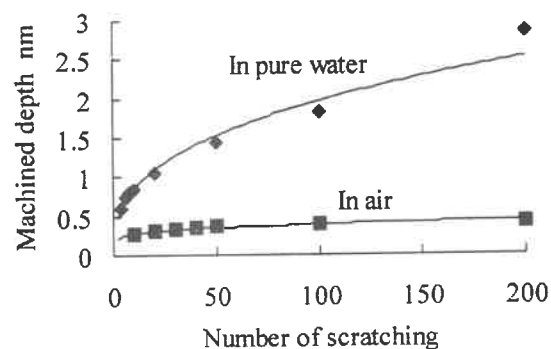


Fig. 9 Relationship between machined depth and number of scratching

4. Influence of surrounding conditions on cutting

4.1 Machined depth

To investigate the influence of surrounding conditions, the workpiece is scratched in air and in pure water under the low wear conditions of the tip.

Fig. 9 shows the relationship between the machined depth and the number of scratching. The same area with the dimensions of $0.5 \times 0.5 \mu\text{m}$ is machined 200 times with scanning speed of $10 \mu\text{m/s}$.

The machined depth in air is smaller than that in pure water. The machined depth in air is saturated after about 50th machining though the machined depth in pure water is still increasing. This tendency is almost the same under any conditions of the initial surface, ambient temperature and humidity.

4.2 Change of surface condition

The surface before and after machining is compared by using the frictional force microscope (FFM) mode.

Fig. 10 shows an AFM image of the surface and an FFM image near the boundary of the machined area after 200th machining in air. The AFM image indicates the machined rectangular area. By considering the machined depth and the result measured with the FFM image, the surface characteristics of the machined area is different from non machined area. Since there are adherent chips and oxide layers on the surface of the machined area, it is not sure that single crystal silicon can be machined with the tip of the AFM in air. The adherent chips on both the tip and the machined surface reduce the removal rate with the tip. In water, the chips hardly affect the sharpness of the tip or chemical reaction assists the removal of silicon.

5. Conclusions

Nanometer cutting has been carried out with an AFM. The following conclusions can be drawn through the experiments:

- (1) The tip wear can be measured by the proposed method during scratch machining. The tip wear is proportional to the traveling distance of the tip in machining.
- (2) The removed amount of silicon with the AFM in water is larger than that in air.

The influence of chemical reaction and other atmosphere will be investigated in future.

Acknowledgment

The authors wish to thank to Prof. S. Shimada of Osaka Univ., Prof. T. Inamura of Nagoya Inst. Technol., and Mr. R. Kokawa and Mr. Y. Yamakage of Shimadzu Corp.

References

- [1] H. Fujita: A Decade of MEMS and Its Future, Proc. 10th IEEE Int. Workshop on Micro Electro Mech. Syst., Nagoya, Japan, pp. 1-8 (1997).
- [2] S. Miyake, M. Ishii, T. Otake, N. Tsushima: Nanometer-scale Mechanical Processing of

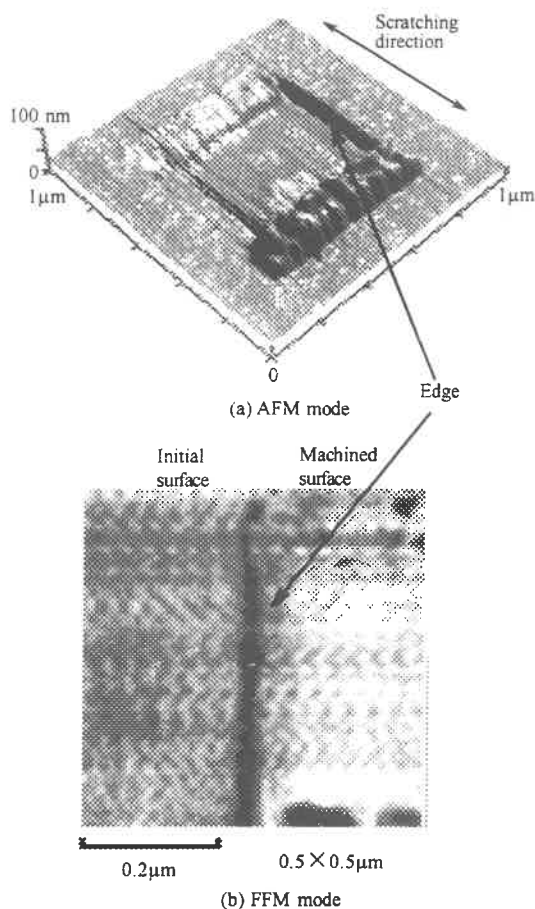


Fig. 10 AFM image of surface and FFM image near boundary of machined area after 200th machining in air

Muscovite Mica by Atomic Force Microscope, J. Jpn. Soc. Prec. Eng., 63, 3, pp. 426-430 (1997) [in Japanese].

- [3] P. Markiewicz, M. C. Goh: Simulation of Atomic Force Microscope Tip-sample/ Sample-tip Reconstruction, J. Vac. Sci. Technol., B, 13, 3, pp. 1115-1118 (1995).
- [4] S. Shimada, N. Ikawa, T. Inamura, N. Takezawa, H. Ohmori, T. Sata: Brittle-Ductile Transition Phenomena in Microindentation and Micromachining, Annals of the CIRP, 44, 1, pp. 523-526 (1995).
- [5] B. Bhushan: Handbook of Micro/ Nano Tribology, CRC Press, (1995)

High-Precision Magnetic Levitation Stage for Photolithography

Won-jong Kim and David L. Trumper

Massachusetts Institute of Technology
Cambridge, MA 02139

1. Introduction

We have designed a permanent-magnet linear motor with Halbach magnet arrays to provide drive force as well as suspension force for a magnetically levitated wafer-stepper stage [1, 6]. Figure 1 shows a perspective view of the selected design concept for prototyping. Figure 2 is a photograph of such a prototype stage designed and implemented by the authors. The platen is levitated without contact by four novel permanent-magnet linear motors which provide both suspension and drive forces. By coordinating the forces applied by each motor, the stage can be accurately controlled in six degrees of freedom. The platen mass of 5.58 kg is supported against gravity by the combined forces of the four motors. Each motor consumes about 5.4 W to lift the platen. The present design has a travel of 50 mm in x and y , a travel of 400 μm in z , and is capable of milliradian-scale rotations about each of these three axes. Detailed fabrication issues of the stage can be found in [4] and are not repeated here. This electromechanical design and configuration has also been adopted in a scanning stage being under testing by Mike Holmes at the University of North Carolina-Charlotte [2].

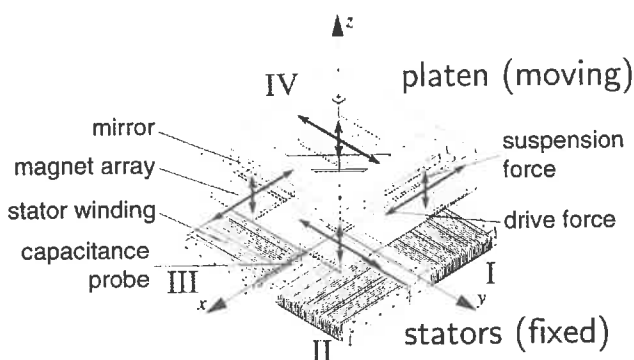


Figure 1: Perspective view of the selected design concept and force definitions of the levitator

This paper presents the working principle underlying the stage operation, and experimental data demonstrating the stage performance capabilities. In the following section, an overview of the magnetic levitation stage is given. Section 3 describes the real-time control of the stage. Test results are given and discussed in Section 4.

2. High-Precision Planar Magnetic Levitator

For our three-phase permanent-magnet linear motors, the parameters have the following values: magnet remanence, $\mu_0 M_0 = 1.29$ T, pitch $l = 25.6$ mm, and the nominal motor air gap $z_0 = 250$ μm . The magnet array thickness is $\Delta = l/4$ and the winding thickness is $\Gamma = l/5$.

To drive the motors, we implemented linear transconductance power amplifiers, with maximum current and voltage ratings of ± 1.5 A at ± 22 V. The nominal power dissipation per motor is 5.4 W at the 0.5-A nominal peak phase current. The suspension power coefficient of the stage is 7.2 mW/N².

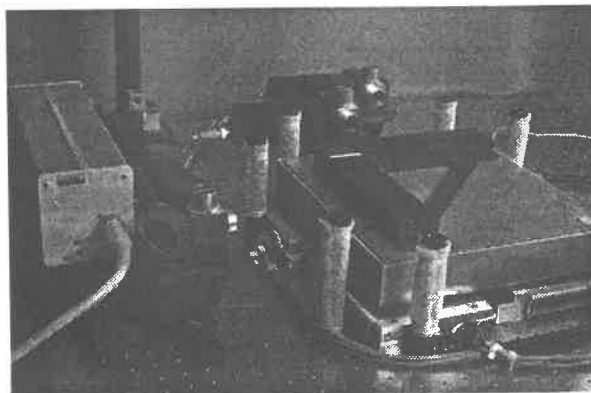


Figure 2: High-precision magnetic levitation stage for photolithography

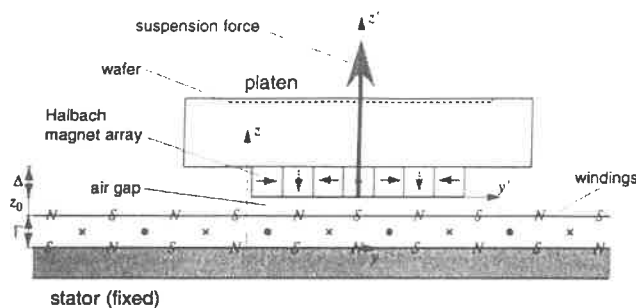


Figure 3: Suspension of the platen in dynamic equilibrium. Only peaks of winding currents are indicated as dots and crosses.

The resistance and the self-inductance of one-phase winding are $14.4\ \Omega$ and $3.44\ \text{mH}$.

Figure 3 depicts the dynamic equilibrium around which the levitator operates and is stabilized. These currents are shown as x's into the page and •'s out of the page. We also show the north (N) and south (S) poles generated by these currents. At a fixed time, if we generate the current sinusoidally distributed in the y -direction as in Figure 3, there is vertical repulsive force between the same magnetic poles of the magnet array and the current distribution (i.e., corresponding north to north and south to south). This vertical repulsive force lifts the platen against gravity. Since this equilibrium is unstable in the lateral direction (in the y -direction in the figure), however, we need active feedback control to stabilize the motion of the platen around this dynamic equilibrium. Conceptually, we can control the magnitude of the force by changing the magnitude of the current, and direction of the force by commutation. See [4, 6, 3] for more details.

The four motors in Figure 1 collaboratively generate all six-degree-of-freedom motions for focusing and alignment and large two-dimensional step and scanning motions for high-precision positioners such as a wafer stepper stage in semiconductor manufacturing. Two of the motors (I and III) drive the stage in the x -direction, and two of the motors (II and IV) drive the stage in the y -direction. The motor forces are coordinated appropriately to control the remaining degrees of freedom.

In the following section, we describe the real-time control of the magnetic levitator.

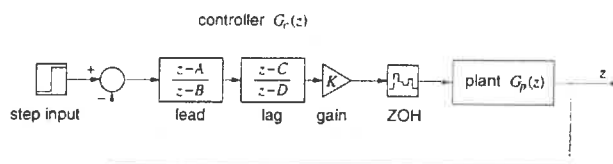


Figure 4: Decoupled lead-lag digital controller

3. Real-Time Digital Control

Figure 4 is a block diagram of the discrete-time lead-lag controller with $A = 0.96300$, $B = 0.68592$, $C = 0.99624$, and $D = 1$. The gains K for the six axes, x , y , z , ψ , θ , and ϕ are $2.2261 \times 10^6\ \text{N/m}$, $2.2261 \times 10^6\ \text{N/m}$, $2.3141 \times 10^6\ \text{N/m}$, $2.2504 \times 10^4\ \text{N/rad}$, $2.2504 \times 10^4\ \text{N/rad}$, and $3.9804 \times 10^4\ \text{N/rad}$, respectively. For small angular motions in our levitator, the Euler angles ψ , θ , and ϕ can be considered as rotational angles around the x -, y -, and z -axes, respectively. Because of the geometrical symmetry in x and y of the levitation system, they have identical controllers. The same is true for the controllers for ψ and θ .

Figure 5 shows a $5\text{-}\mu\text{m}$ step response in x with this lead-lag compensator. All the six-axis controllers are operational in this experiment. Note that the platen dynamics are coupled in all six degrees of freedom. Because of this coupling nature of the platen, there are perturbed motions in all other five axes. The step response also shows a 40-Hz frequency component which we believe to be a natural mode of the closed-loop system. The low damping shows that the system has only about 20° phase margin as compared with a design value 45° , and this response show correspondingly more ringing compared with the computer simulation. An unmodeled coupling effect is likely responsible for this phenomenon. The source of this coupling remains to be determined, and will be investigated in future work.

4. Test Results

We provide several test results in this section. Figure 6 shows the position regulation results by using the decoupled lead-lag controllers designed above. The platen's position is held at the origin in the global coordinate frame. The position noises in x - and y -axes are on the order of $5\ \text{nm rms}$. The

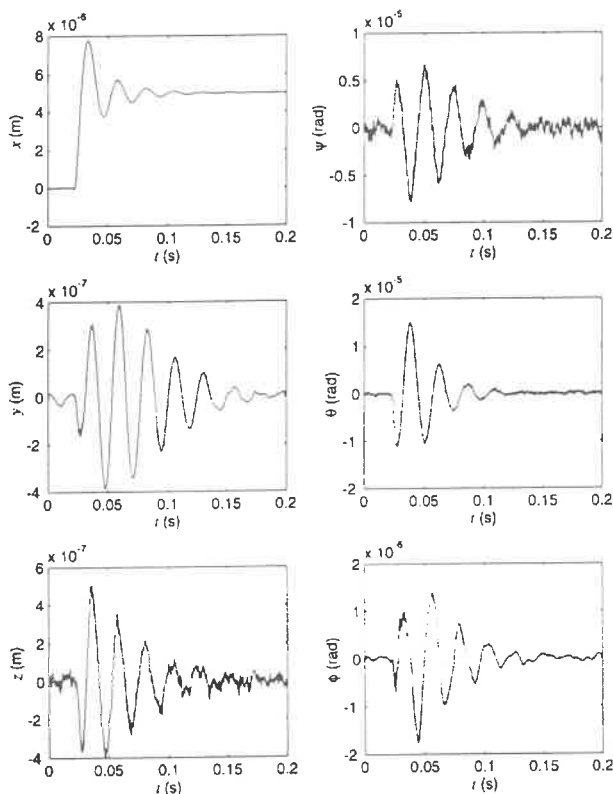


Figure 5: 5- μm step response in x with perturbed motions in the other five axes

position noise in ϕ is 0.05 μrad rms. The vertical displacement z shows a 70-nm-order position error envelope. The A/D converter electronics noise is primarily responsible for the poorer position noise in z , relative to x and y which are measured with sub-nanometer resolution via our laser interferometer system.

The testing result in Figure 6 also has a dominant 120-Hz noise. Lab floor vibration due to an HVAC equipment located in the next room is believed to generate this noise. The floor vibration of our lab includes other lower frequency components at 7.5 Hz, 8 Hz, 15 Hz, and 28 Hz. A more detailed trace of this floor vibration can be found in [5]. This vibration is attenuated by an air spring isolation system, but is still dominated in the error motions.

For the fastest motion generation, we accelerate the platen at the highest acceleration possible until the velocity reaches the maximum slew rate, hold this velocity for a time, then decelerate the platen at the same maximum acceleration to brake the platen

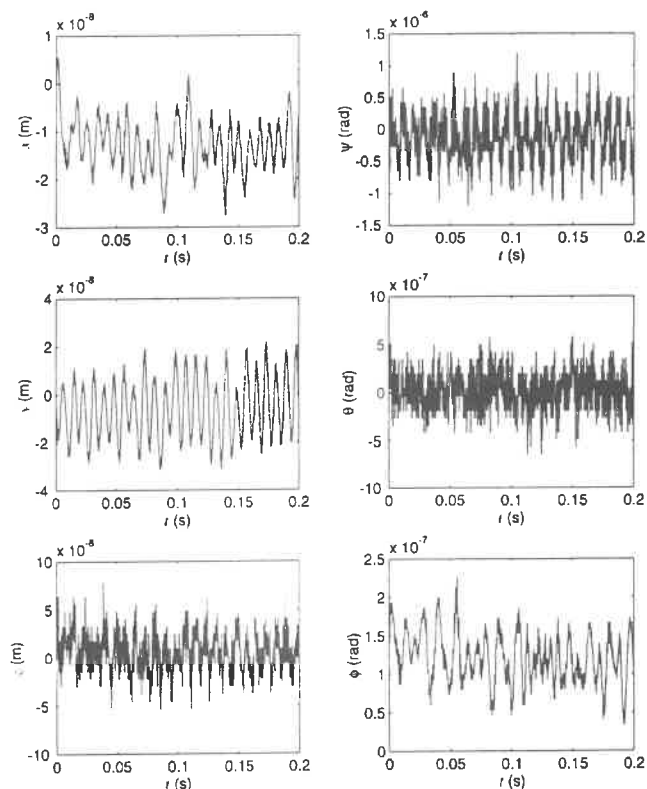


Figure 6: Position noises in six-degree-of-freedom position regulation with the six decoupled lead-lag compensators

motion. This is called a trapezoidal velocity profile. The platen thus follows a parabolic, a linear, and another parabolic reference trajectories. Figure 7 is a test result with a 20-mm step in the y -direction. The acceleration given to the platen is 10 m/s^2 (about 1 g). The command follows a 120-ms transition time for the 20-mm step.

We have also demonstrated that the stage can simulate motions of a wafer stepper. The platen makes step-and-settle motions described above in the positive y -direction, steps over in the positive x -direction, makes step-and-settle motions in the negative y -direction, and so on. The platen can also follow a circle with 30-mm diameter in one second. The purpose of this demonstration is to show the levitation system can generate any unified two-dimensional motions with predetermined paths.

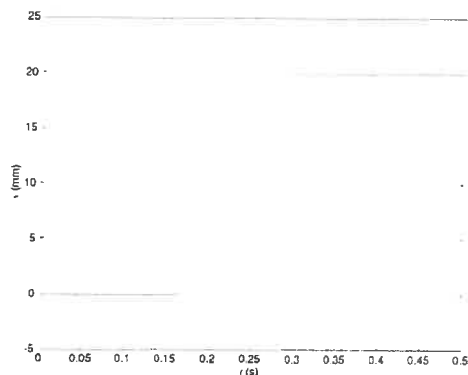


Figure 7: 20-mm step in y . The platen follows a trapezoidal velocity profile.

5. Conclusions

We have designed and implemented a world's first high-precision six-degree-of-freedom magnetic levitator with large two-dimensional motion capability for photolithography in semiconductor manufacturing. The magnetically levitated platen generates all the required small motions for focusing and alignments as well as large planar motions for wafer positioning. This magnetic levitation stage can be readily used in a clean room, since there is no wear particle generation and no lubrication is required. This design is also highly suitable for vacuum environments, since the heat generated in the motors is in the fixed frame and thus can be readily removed by material conduction.

Furthermore, there is no backlash because the levitation system uses no intermediate power transmission device like ball screws. Thanks to the lack of friction or backlash, the position accuracy depends primarily on sensors. Thus, the stage precision depends primarily on the fundamental limits of metrology and control. Another advantage is that no fine finishing of surfaces of mechanical parts, such as for bearings, is necessary. This simplifies the production process and reduces the manufacturing cost.

The magnetic levitation stage has been tested successfully. We implemented decoupled lead-lag digital controllers in a 320C40 digital signal processor. The sampling rate of the system is 5 kHz. The control loop for the levitator is able to be closed at a 50-Hz bandwidth. A mirror resonance at 400 Hz limits the ability to achieve higher bandwidth. Important

experimental achievements include 5-nm rms position error in x and y , 30-nm rms position error in z , 20-mm steps following 120-ms references, and 1- g acceleration. We thus have demonstrated that the planar magnetic levitator is a promising candidate as the positioning stage in next-generation semiconductor manufacturing.

Acknowledgments

Portions of this work are part of a thesis submitted by W.-J. Kim [3] for the Doctor of Philosophy degree in Electrical Engineering and Computer Science at the Massachusetts Institute of Technology, Cambridge, Massachusetts. This work is supported in part by a contract from Sandia National Laboratories under Subcontract No. AH-4243 and in part by a Scholarship from Korean Ministry of Education to W.-J. Kim and a National Science Foundation Presidential Young Investigator Award (DDM-9496102) to D. L. Trumper.

References

- [1] K. Halbach. Design of permanent multipole magnets with oriented rare earth cobalt material. *Nuclear Instruments and Methods*, 169(1):1-10, 1980.
- [2] M. L. Holmes, R. J. Hocken, and D. L. Trumper. A long range scanning stage (the LORS project). In *Proceedings of 1997 ASPE Annual Meeting*, October 1997.
- [3] W.-J. Kim. *High-Precision Planar Magnetic Levitation*. PhD thesis, Massachusetts Institute of Technology, June 1997.
- [4] W.-J. Kim, D. L. Trumper, and J. B. Bryan. Linear-motor-levitated stage for photolithography. *Annals of the CIRP*, 46(1):447-450, August 1997.
- [5] S. J. Ludwick. Modeling and control of a six degree of freedom magnetic/fluidic motion control stage. Master's thesis, Massachusetts Institute of Technology, February 1996.
- [6] D. L. Trumper, W.-J. Kim, and M. E. Williams. Design and analysis framework for linear permanent-magnet machines. *IEEE Transactions on Industry Applications*, 32(2):371-379, March/April 1996.

A Long-Range Scanning Stage (The LORS Project)

Mike Holmes, Robert Hocken

Center for Precision Metrology, University of North Carolina,
Charlotte, NC, USA

David Trumper

Laboratory for Manufacturing and Productivity, Massachusetts Institute of Technology,
Cambridge, MA, USA

This extended abstract describes a magnetically-suspended six-degree-of-freedom precision motion control stage. This Long-Range Scanning (LORS) stage has travel in the horizontal plane of 25 mm x 25 mm along with 100 μ m of vertical travel. Vertical position feedback is provided by three capacitance probe sensors while heterodyne laser interferometry is used for lateral position feedback. The performance objectives are a positioning resolution of 0.1nm, positioning repeatability of 1nm, and a positioning accuracy of 10nm. These performance objectives have been chosen to match the measurement requirements associated with present and future production needs for devices such as integrated circuits, photo-masks, and micro-mechanical actuators. The stage is presently operational in a preliminary fashion; full operation is expected shortly.

Keywords: magnetically suspended; precision motion control

Project Overview

The purpose of this research is to explore the use of magnetic bearings in precision motion control applications. In previous research^{1,2}, we demonstrated the utility of combining magnetic bearings with neutrally-buoyant oil flotation to achieve sub-nanometer positioning resolution. This project furthers those efforts by increasing the range of travel in the horizontal plane to 25 mm² and by placing an ambitious accuracy specification which if met, will make this stage one of the world's most accurate positioning stages. Scanned probe microscopy will be used to characterize the performance of the LORS stage. A two dimensional grid of crosses has been fabricated for use as a reference artifact and will be referred to hereafter as the **sample grid**. The sample grid was manufactured by Hewlett-Packard³ and measured at PTB⁴. Our goal is to also measure the sample grid and to compare our results with those of PTB. Later the stage will be error mapped using techniques similar to those described by Ye et al⁵. Herein we summarize the results to date of this 2 $\frac{1}{2}$ year old project and present some photographs of the experimental setup.

The LORS Stage

The LORS stage, shown in figure 1, consists of a moving element (**platen**), a **machine frame**, and a **metrology head**. The bottom of the platen holds four permanent magnet arrays. A **reference block** is mounted to the top of the platen via three flexures⁶. The platen is buoyantly floated in oil to reduce bias currents in the motors and to provide

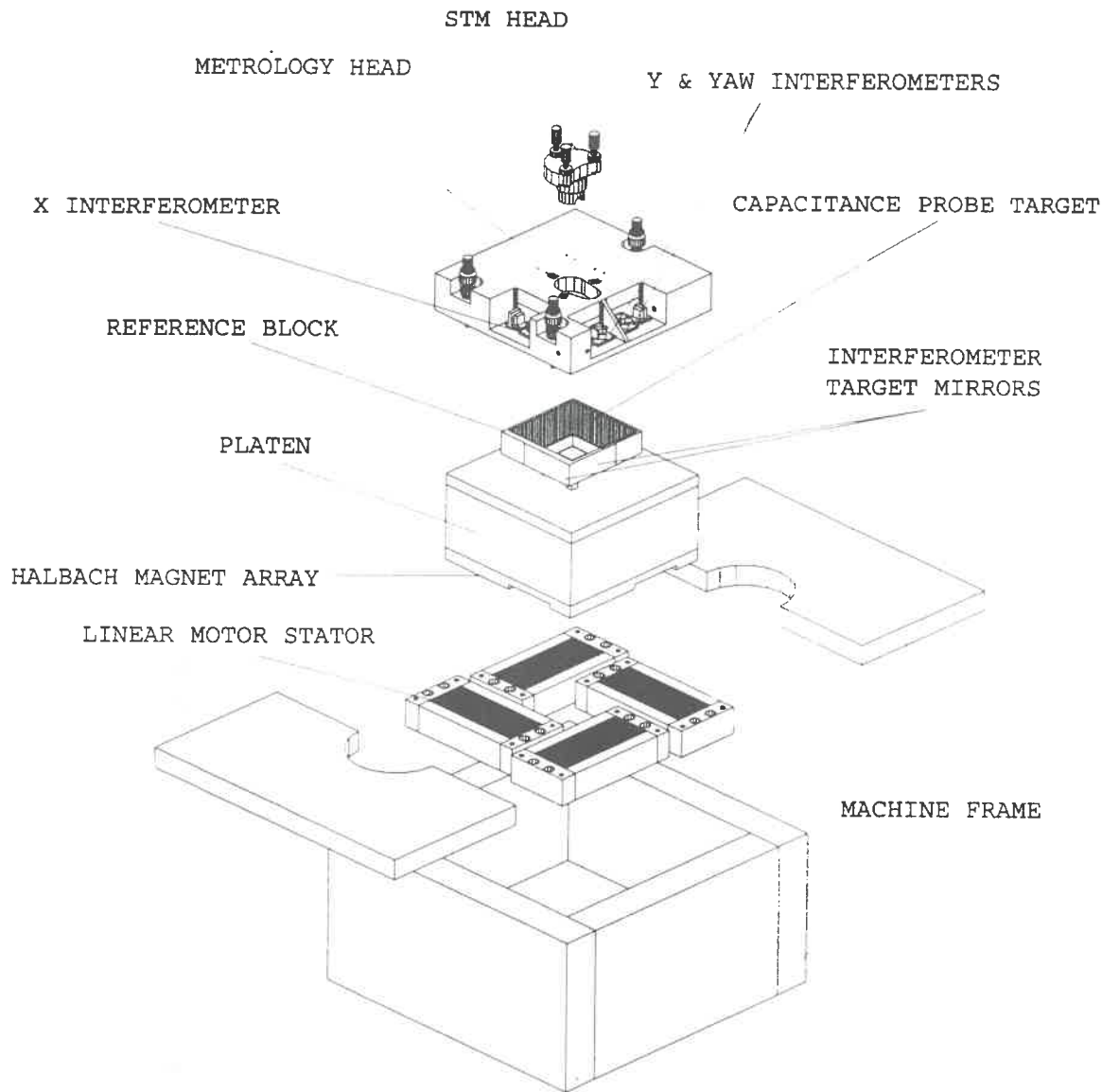


Figure 1: The LORS Stage (Exploded View)

viscous damping and high-frequency coupling. The stator windings are mounted in the machine frame opposite the magnet arrays. The reference block contains the reference surfaces for the position feedback sensors, namely, target mirrors for the interferometers and a conductor plane which acts as the target(s) for the capacitance probes. The sample to be imaged is supported by a sample holder which is kinematically mounted to the reference block. The metrology head is kinematically mounted to the top of the machine frame. The metrology head contains the feedback sensors, three 8-pass interferometers for lateral position feedback and three capacitance probes for vertical position feedback. The metrology head also provides support for the scanned probe microscope head. To reduce the effects of thermal drift, the metrology head and the reference block are made of Zerodur, with a coefficient of thermal expansion of $0.05 \mu\text{m}/\text{m}/^\circ\text{C}$. The idea is to control the currents in the motors, thereby controlling the forces exerted on the platen, based on

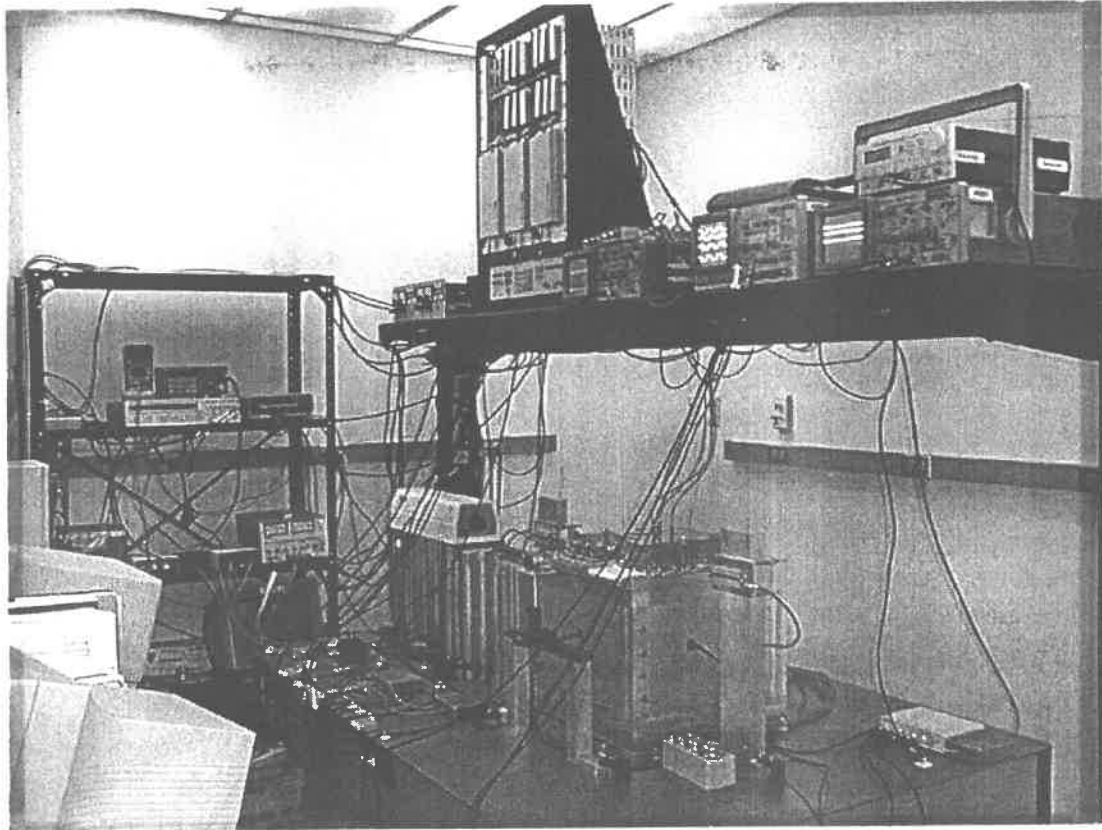


Figure 2: The LORS Project Support Hardware

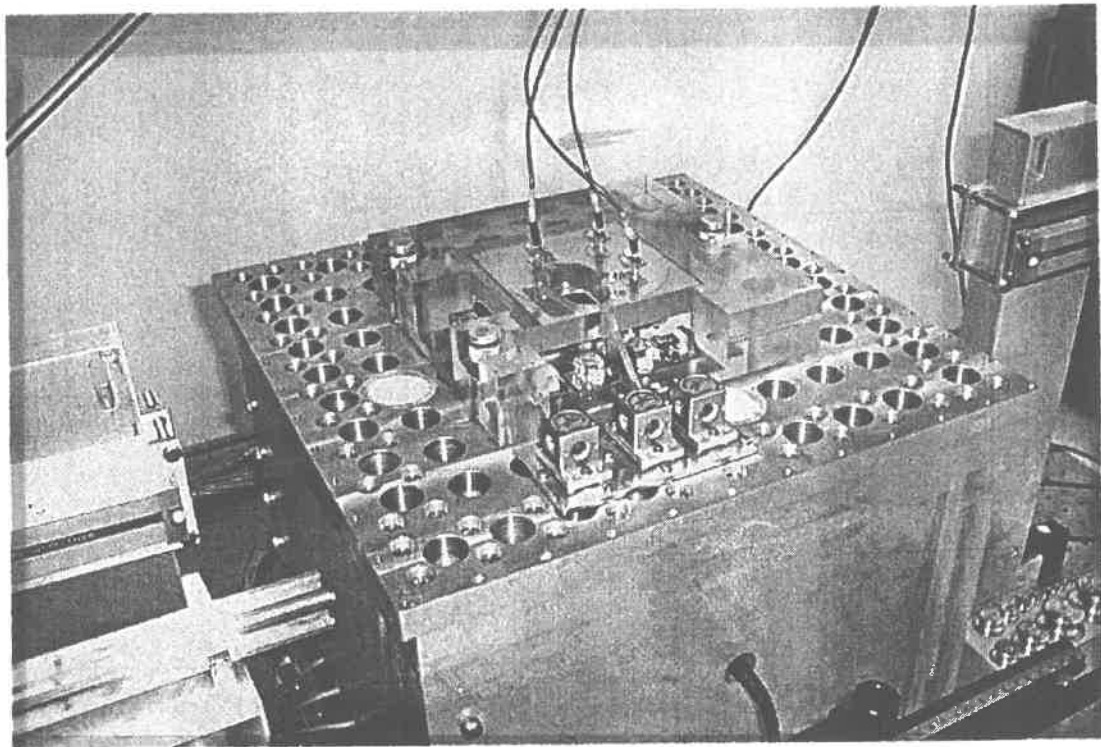


Figure 3: The LORS Stage

the position feedback from the interferometers and the capacitance probes. We will be able to image various samples which will be placed on the sample holder. The scanned probe interface mount is designed to support various types of scanned probe microscope heads (i.e., Atomic Force Microscopes (AFM) and Scanning Tunneling Microscopes (STM)).

Acknowledgments

This research is funded by the National Science Foundation under grant DMI-9414778. The ADE Corporation⁸ is providing the capacitive gaging electronics and Zygo Corporation⁹ is assisting with the metrology head and reference block fabrication. Hewlett-Packard Photomask fabricated the sample grid which will ultimately be used for performance characterization. The relative locations of the crosses on the sample grid have been measured by PTB. It is our intent to use their results as a comparison rule. Dr. Gerald Blessing of the National Institute of Standards and Technology did some initial experiments to determine the feasibility of using ultra-sonic probes to measure the initial position of the platen in the silicone oil. Also, Won-jong Kim at the Massachusetts Institute of Technology developed the theory underlying the linear motor design^{10,11} and provided the impetus for using the chosen coil configuration.

References

- [1] Holmes, M.L, Trumper, D.L., and Hocken, R. J., "Atomic-scale precision motion control stage (The Angstrom Stage)," *CIRP Ann.*, vol. 44, no. 1, pp.455-460, 1995.
- [2] Holmes, M.L., and Trumper, D.L., "Magnetic/fluid-bearing stage for atomic-scale motion control (the Angstrom stage)," *Precision Eng.*, vol. 18, no. 1, pp. 38-49, Jan. 1996.
- [3] Hewlett-Packard Photomask, 5301 Stevens Creek Blvd., Santa Clara, CA 95052-8059.
- [4] Physikalisch Technische Bundesanstalt (PTB), Braunschweig, Germany.
- [5] Ye, J., Takac, M., Berglund, C., Owen, G., and Pease, R., "An exact algorithm for self-calibration of two-dimensional precision metrology stages," *Precision Eng.*, vol. 20, no. 1, pp. 16-32, Jan. 1997.
- [6] Smith, S.T., and Chetwynd, D.G., *FOUNDATIONS OF ULTRAPRECISION MECHANISM DESIGN*, Gordon and Breach Science Publishers, Yverdon, Switzerland, 1994.
- [7] Slocum, A., *Precision Machine Design*, Prentice Hall, Englewoods Cliffs, NJ, 1992.
- [8] ADE Technologies, 77 Rowe Street, Newton, Mass. 02166.
- [9] Zygo Corporation, Laurel Brook Road, Middlefield, Conn. 06455.
- [10] Trumper, D.L., Kim, W.J., and Williams, M.E., "Design and Analysis Framework for Linear Permanent-Magnet Machines," in press, *IEEE Trans. Ind. Applic.*, vol. 32, no. 2, Mar/Apr 1996.
- [11] Kim, W.J., "High-Precision Planar Magnetic Levitation," Ph.D. Thesis, Electrical Engineering and Computer Science, Massachusetts Institute of Technology, June 1997.

A 3 DOF Z-tilts Micro Positioning System Using Electromagnetic Actuators With Air Bearings

Sung-Q Lee and Dae-Gab Gweon

Dept. of Mechanical Engineering

Korea Advanced Institute of Science and Technology

373-1 Kusung-Dong, Yusong-Gu, Taejon, 305-701, Korea

Abstract

A new 3 DOF (Z, pitch and roll motion) micro positioning system is developed. It uses electromagnetic actuator and air bearings. An Electromagnetic actuator can produce attraction force continuously and air bearings are used as the suspension and guidance with the clearance of several ten- μm in z-direction. Usually, active magnetic bearings with electromagnetic actuator can guide one degree of freedom and magnetic flux is very difficult to analysis. However, air bearing floating on the surface can guide three degrees of freedom, $XY\theta$, moreover, can provide the counter force to the electromagnetic attraction force to make the z-directional positioning without affecting the magnetic flux of the electromagnet. With the control of current, equilibrium position between magnetic attraction force and air bearing thrust force can be controlled with inherently infinite resolutions. With the open loop control, this system provides 30 nm positioning resolution over the total range of 50 μm along the z-direction and 0.6 arcsec resolution over the total range of 100 arcsec in pitch and roll motion.

1. Introduction

A new 3 DOF (Z, pitch and roll motion) micro positioning system is developed. It uses electromagnetic actuator and air bearings. Electromagnetic actuator can produce attraction force continuously so that the resolution depends on measurement devices and D/A convert. Air bearings are used as the suspension and guidance with the clearance of several ten- μm in z-direction[1][2][3]. Usually, active magnetic bearings with electromagnetic actuator can guide

one degree of freedom and magnetic flux is very difficult to analysis. However, air bearing floating on the surface can guide three degrees of freedom, $XY\theta$, moreover, can provide the counter force to the electromagnetic attraction force to make the z-directional positioning without affecting the magnetic flux of the electromagnet[4][5].

The principle of this system is similar to the magnetic levitation system except it uses air bearings as suspension. Therefore, this system is

applicable to the magnetic levitation applications, wafer positioning system and another measurement systems which need fine Z-tilt motions.

2. System Configuration and Principle

This system is suspended by three air bearings which are radially disposed with 120 degrees pitch, about 57 mm radius and are wound by coils to play the role of the electromagnetic actuator. From the combination of three actuators, Z-tilt motions can be obtained. (figure 1.)

Electromagnetic Parts

As in Figure2, magnetic flux is induced by the coils and closed along the base plate, air bearing and sensor holder. An Electromagnetic attraction force is induced between the base plate and sensor holder (h_o) or between the base plate and air bearing pad (h_g). However, since the h_g is several ten- μm which is much smaller than h_o , the attraction force is largely induced in the gap of h_o . Therefore attraction force is[1],

$$F_{\text{attrac.}} = (N I)^2 \mu_o A / (2 h_o)^2$$

where N is the number of coil turns, I, current, μ_o , permeability of free space, A, the cross section of effective area, h_o , the gap between the base plate and sensor holder, respectively. Within the Z-directional range, for h_o varies slightly, attraction force is only the function of current I.

Air bearing Parts

Two main performances of air bearings are bearing stiffness and load capacity. Design parameters to decide these performances are supply pressure, orifice diameter, the number of orifices, air gap size and the size of bearing. In

this system, annular type air pads are used. To play the role of z-directional counter force to electromagnetic force, it is designed to have load capacity of 9kgf at $10\mu\text{m}$ h_o , 1kgf at $50\mu\text{m}$ with 4 orifices of 0.3 mm diameter. Since it is well known to design air bearings[1], it isn't difficult to design exactly. Theoretically load capacity has nonlinear characteristics versus h_g as shown in figure 3.[1] Accordingly, since the stiffness is not constant, open loop control may bring different errors at each point.

Principle of 3 DOF motion

An Electromagnet exerts 0-20kgf according to current I (0 - 2 Ampere) and air bearing's load capacity is 9kgf at $10\mu\text{m}$ h_o , 1kgf at $50\mu\text{m}$ as mentioned above. Therefore, with the control of current, equilibrium position between magnetic attraction force and air bearing thrust force can be controlled with inherently infinite resolutions. From the combination of three actuators, Z-tilt motions can be obtained. To decouple the coupling of three actuators, hinge mechanism is used(figure2).

3. Experimental Results

Figure 4 shows a schematic of the Z-tilt stage control. The overall system consists of a PC, D/A converters, current drivers, displacement measurement sensors, A/D converters and the system structure. The PC provides command input voltage (0-10 V) via the home-built D/A converters(12bit). The voltage is amplified by the current drivers (0-2 Ampere) and fed to the electromagnetic actuators. The actuation provides

the displacement required for the system positioning. The actual position information detected by the capacitance gages (probe 2805 with the sensitivity of $25\mu\text{m}/10\text{V}$, ADE corp.) is also fed back to the PC via the 16 bit A/D converter (20ksamples/sec, AT-MIO-16XE50, National Instruments).

Figure 5 shows the experimental results of open loop control of one actuator remaining another two actuators at constant position. As in the figures, the system is commanded to move multi-step of $2\mu\text{m}$. As can be seen from the figure, the system has noticeable open loop errors due to the effect of non-uniformity of bearing stiffness, and a small time varying fluctuating. Thus closed loop control must be performed. Though it uses hinge mechanism to decouple the each actuation motion, there are some coupling effects at each step. To obtain the resolution, it is commanded to move smaller step until distinguishable level. It is noticeable that resolutions are level of sensor noise. Therefore the resolution depends on sensor to a certain degree. With the open loop control, this system provides 30 nm positioning resolution over the total range of $50\mu\text{m}$ along the z-direction (figure 6) and accordingly, 0.6 arcsec resolution over the total range of 103 arcsec in pitch and roll motion.

4. Conclusion

A new 3 DOF (Z, pitch and roll motion) micro positioning system is developed. It uses electromagnetic actuator and air bearings. An Electromagnetic actuator can produce the attraction force continuously 0-20kfg and air

bearings are used as the suspension and guidance with the clearance of several ten- μm in z-direction. With the control of current, equilibrium position between magnetic attraction force and air bearing thrust force can be controlled with inherently infinite resolutions.

With the open loop control, this system provides 30 nm positioning resolution over the total range of $50\mu\text{m}$ along the z-direction and 0.6 arcsec resolution over the total range of 103 arcsec in pitch and roll motion.

References

- [1] Alexander H. Slocum, 1992, "Precision Machine Design", Prentice-Hall, pp551-628.
- [2] S.A.Nasar and L.E.Unnewehr, 1983, "Electromechanics and Electric Machines", John Wiley and Sons, pp124-134.
- [3] Francis. C. Moon, 1984, "Magneto-Solid Mechanics", John Wiley and Sons, pp275-279.
- [4] Shigeo KATO, 1996, "Theoretical analysis of Damping Characteristics of Air-Lubricated Slideways", Int. J. JSPE, Vol. 30, No. 1, pp77-82.
- [5] Yoshiyuki Tomita, Yasushi Koyanagawa and Fumiaki Satoh. 1996, "A Surface Motor-driven Precise Positioning System", Precision Engineering, Vol.16, No.3, pp184-191.

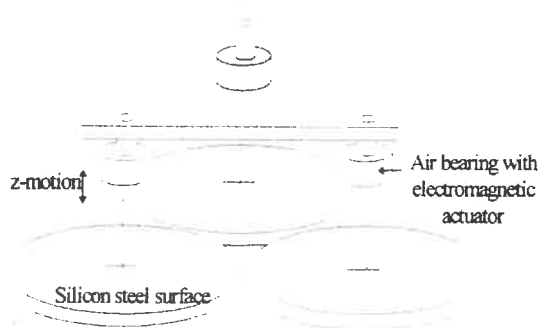


Fig. 1. The schematic of Z-tilts micro positioning system

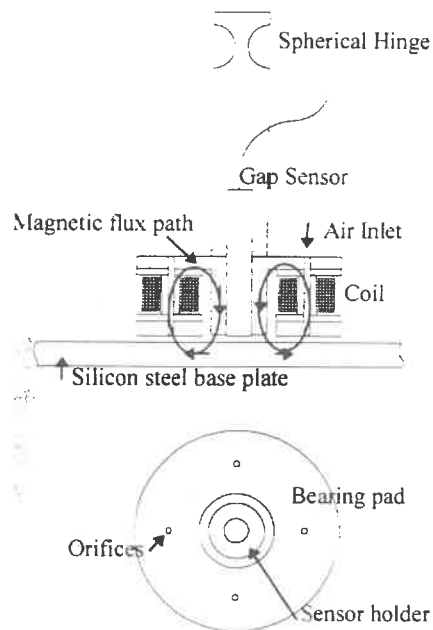


Fig. 2 Electromagnetic Actuator with Air Bearing

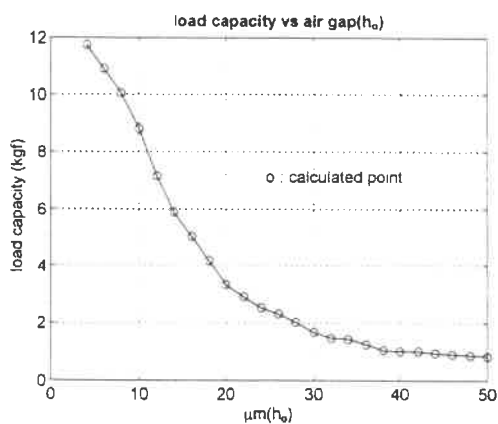


Fig. 3 Bearing's load capacity vs air gap(h_o)

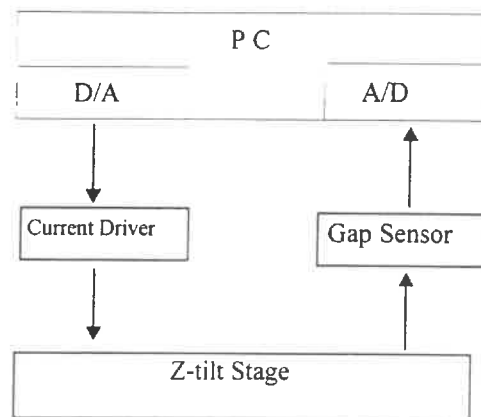


Fig. 4 Schematic of Experimental system

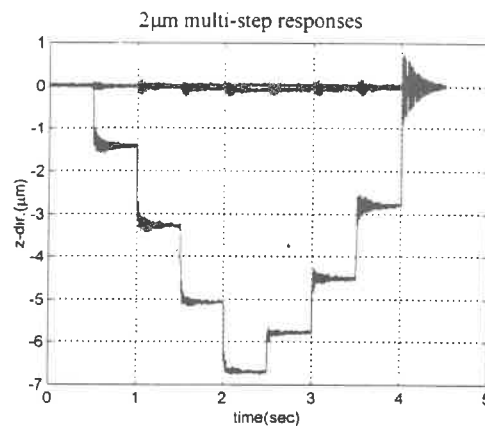


Fig. 5 Open loop multi-step responses

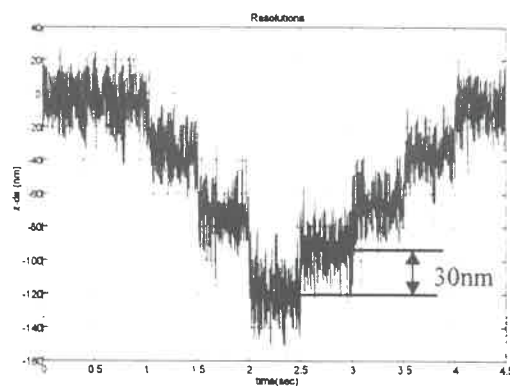


Fig. 6 Resolution of Z-tilt stage

A 1/100 Machine Tool to Fabricate 1/100 Micro Parts

Kazuhisa ISHII, Masayuki NAKAO and Yotaro HATAMURA
The University of Tokyo, Faculty of Eng., Dept. of Engineering Synthesis
7-3-1 Hongo, Bunkyo-ku, Tokyo 113 Japan

1. Introduction

"There's plenty of room at the bottom", in 1959, R. P. Feynman suggested plenty of possibilities for 3-dimensional miniaturized machine [1]. A micro mother machine is capable of producing smaller child machines. Currently, micromachining of micro parts mainly employs lithography technique. The technique has improved for semiconductor processing ; Dr. Feynman paid \$ 1,000 for writing a page with 100 nm high characters using an electron beam exposure machine [2]. But this technique is for only thin-film deposited, 2-dimensional, plate-like structures, not for 3-dimensional ones. We want to fabricate complicated, 3-dimensional, micro structures using various materials, which is the same as conventional macro structures.

The authors challenged to design, construct, and evaluate a "1/100 machine tool," i.e. a machine tool which has by 1/100 a smaller size than conventional machines, to fabricate 1/100 micro parts. Through the evaluations we want to know : design constraints of the 1/100 machine tool, new mechanisms to be developed in future, new market using the machine tool and advantages of the miniaturized machines. Miniaturized machine tools can save the machine cost and facility space ; they can realize smaller positioning error induced by thermal expansion ; they can install in the special environment such as in vacuum or in clean air.

We need to develop some new mechanisms to realize the 1/100 machine tool ; all parts of the 1/100 machine tool could not be miniaturized by 1/100. For example, a voice coil motor mechanism is developed for a table drive though a ball screw and AC motor is used in a conventional machine tool. Information on the development of such partial mechanisms, in addition the synthesized 1/100 machine tool, will be useful for future design of various micro elements.

2. Design of the 1/100 machine tool

Design requirement of the machine is "to realize a 1/100 sized machining center". The miniaturized machining center should have the same functions as conventional macro machining center : 3-axis translating table, spindle for tool rotation, auto-tool-changer and so on.

To satisfy the requirement, we should realize the following fourteen functions. The functions are required for machining centers regardless of size. But their realizations were difficult in miniature world, and had worked as design constraints for us. Next we decide the mechanism individually for fourteen functions as follows.

- (1) table drive : by a voice coil motor
 - (2) table guide : by a parallel plates spring
 - (3) table positioning control : feedback controlled by deformation of the parallel plates spring using strain gauges attached on the plates
 - (4) spindle drive : by an air turbine using high pressured air
 - (5) spindle bearing : by a hydrodynamic bearing
 - (6) origin set-up : adjusted table position observing the workpiece and the tool in upper direction and front-side direction with two microscopic CCD cameras
 - (7) cutting tool : sharpened on a cutting edge to $0.1\text{ }\mu\text{m}$ using a $\phi 300\text{ }\mu\text{m}$, purchased single crystal diamond endmill **, or using a $\phi 50\text{-}100\text{ }\mu\text{m}$, homemade carbide endmill fabricated by Wire Electro-Discharged Grinding [3].
 - (8) tool change : * changing a rotor and a tool together which are glued, or ** by a micro chuck in the spindle
 - (9) chip removal : by air flow exhausted from a $\phi 100\text{ }\mu\text{m}$ glass pipe
 - (10) coolant flow : by air flow exhausted from a $\phi 100\text{ }\mu\text{m}$ glass pipe
 - (11) cutting force measurement : by a force sensor using strain gauges
 - (12) workpiece material : made of plastic materials or metals
 - (13) workpiece chuck : * glued on the plate, which is fixed on the table, or ** by a vacuum chuck
 - (14) workpiece transfer : ** supplied on a hoop sheet
- * tentatively developed mechanism, ** still developing mechanism

3. Fabrication of the 1/100 machine tool

Fig. 1 shows the prototype of the 1/100 machine tool, which is $26\times 26\times 67\text{ mm}$ in size, 167 g in mass. The spindle is $\phi 12\times 16\text{ mm}$ in size, rotating at 138,000 r.p.m. maximum with a $0.1\text{ }\mu\text{m}$ radial vibration. The table

is $20 \times 20 \times 40$ mm in size, and has 3 axes (x, y and z) translational linear drives with a 1 mm working range in x and y axis, and a 0.8 mm in z. The driving force of the table is 1 N maximum, and the rotating torque of the spindle is 0.1 mNm maximum.

Fig. 2 shows the homemade carbide endmill sized $\phi 50 \mu\text{m} \times 200 \mu\text{m}$ long with a $5 \mu\text{m}$ edge roundness. It was made by WEDG in rotating the tool. As applying voltage between the tool and the wire has not optimized yet, the tool surface looks relatively rough. The tool was glued with the rotor of the spindle in tool making : the axis of the tool is coincided with the axis of the spindle rotation. It has a half circle shape in cross section.

4. Evaluation of the 1/100 machine tool

4.1 Evaluation of the mechanisms

Fig. 3 shows a resolution in x axis of the table. The figure indicates that steps of 80 nm are represented. The resolution should be 50 nm.

Fig. 4 shows a force (top) and force-induced displacement (bottom) under no control (left) or under position feedback control to keep an ordered position (right). When the force of about 1.1 mN worked the table, the table moved by $9 \mu\text{m}$ under no control, but by $0.5 \mu\text{m}$ under position feedback control. The control improved the rigidity from 120 N/m to 2,200 N/m.

4.2 Evaluation of the machine in cutting

Fig. 5 shows an optical view in cutting : 60,000 r.p.m. in spindle rotation, $30 \mu\text{m}/\text{sec}$ in feed rate, $30 \mu\text{m}$ in cutting depth and PPS (Poly-Phenylene Sulfide) for the workpiece material. Chips were not flown by air ; melted materials were stuck to the tool.

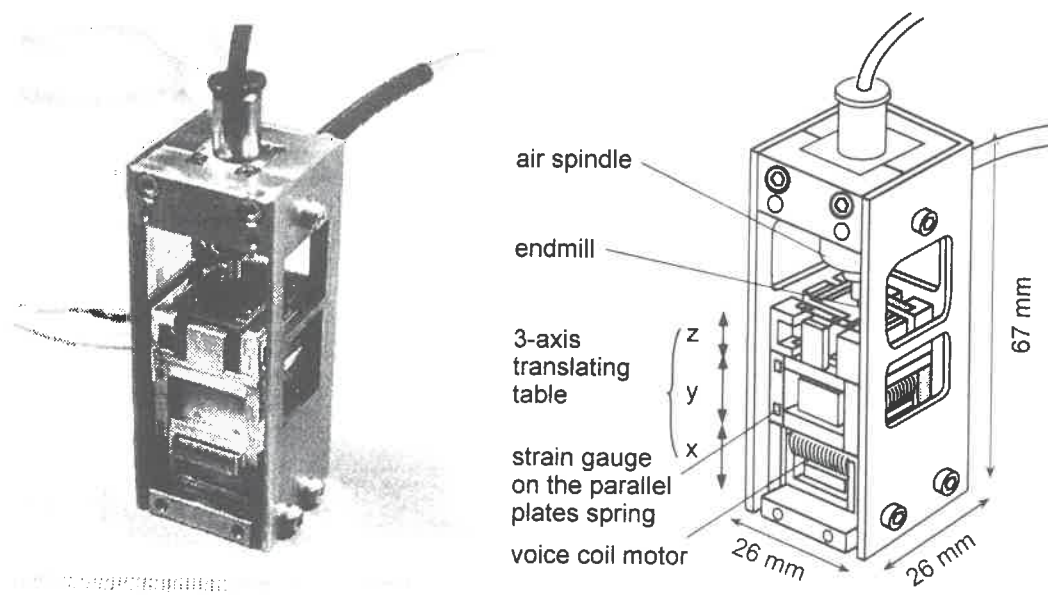


Fig.1 View of micro milling machine

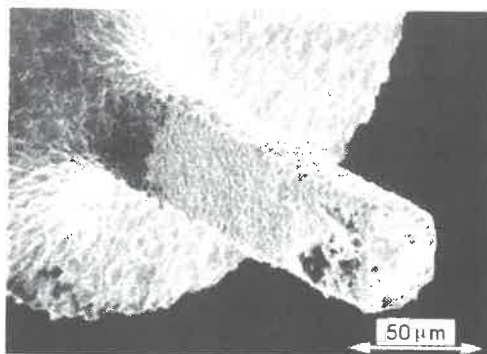


Fig.2 SEM view of carbide endmill

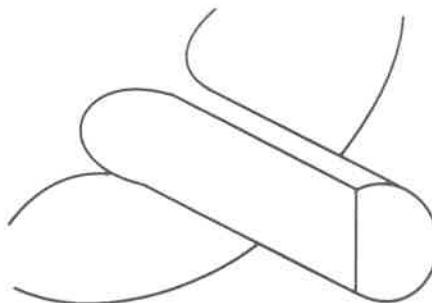


Fig. 6 shows a grooved workpiece using carbide endmill made by WEDG: 60,000 r.p.m. in spindle rotation, 60 $\mu\text{m}/\text{sec}$ in feed rate, 70 μm in groove width, 10 μm in groove depth and PMMA for the workpiece material. The tool was suddenly broken in cutting suddenly at the fixed finish point of the groove; Increased cutting force should break the endmill.

Fig. 7 shows another grooved workpiece using a diamond endmill: 60,000 r.p.m. in spindle rotation, 600 $\mu\text{m}/\text{sec}$ in feed rate, 380 μm in groove width, 30 μm in groove depth and PMMA for the workpiece material. Edge of the groove is free from the chips.

Fig. 8 shows the cutting force measured by a force sensor equipped between the workpiece and the table using (a) carbide endmill and (b) diamond endmill. Fig. 8 (a) indicates that the first slope was due to sensor deformation, the last vibration was due to air blow for chip removal, and the cutting force in stably cutting was about 1 mN. Fig. 8 (b) indicated that the cutting force was less than the resolution of the force sensor even though diameter of the endmill is larger than that of carbide endmill.

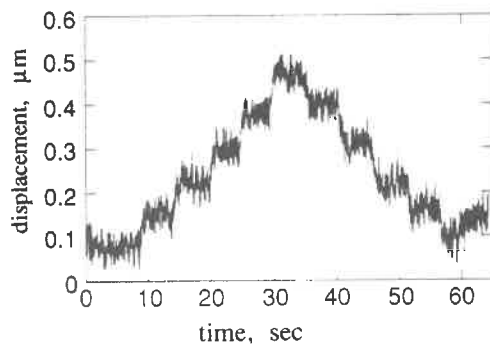


Fig. 3 Step movement in X axis of the table to measure a resolution

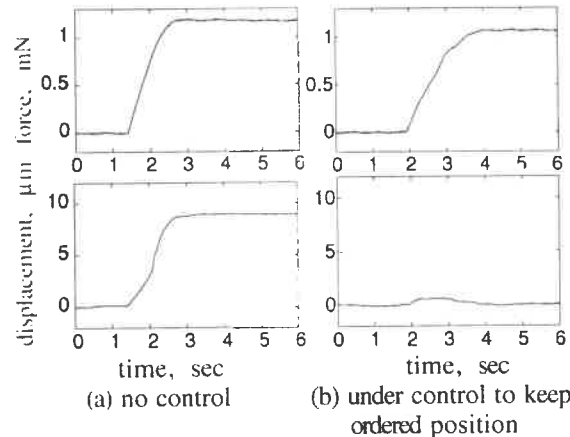


Fig. 4 Force-induced table displacement in X axis of the table

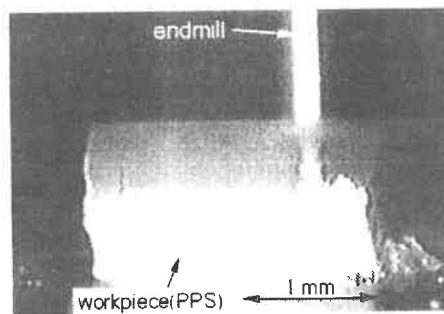


Fig. 5 View of milling process

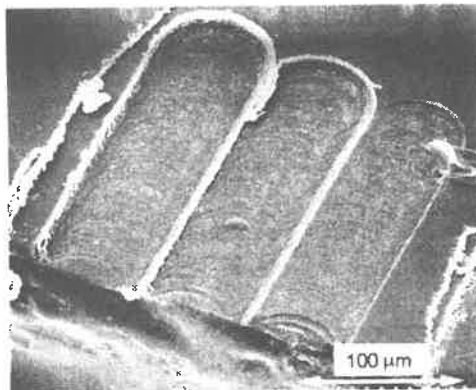


Fig. 6 SEM view of milled workpiece
Cutting condition: rotation: 60 krpm, feed rate: 60 $\mu\text{m}/\text{sec}$
depth: 10 μm , width: 70 μm , workpiece: PMMA

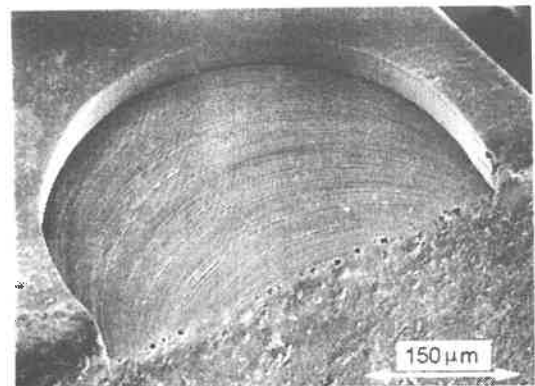


Fig. 7 SEM view of milled workpiece
Cutting condition: rotation: 60 krpm, feed rate: 600 $\mu\text{m}/\text{sec}$
depth: 30 μm , width: 380 μm , workpiece: PMMA

5. Discussion

We could not realize the four mechanisms by ourselves so far : (7) cutting tool (sharp cutting ridge), (8) tool change, (13) workpiece chuck and (14) workpiece transfer.

Sharp cutting ridge is indispensable for micro cutting. Because the tool cuts the workpiece by very low removal rate such as $0.06 \mu\text{m}$ per blade rotation and a dull edge makes only burnishing, not cutting. The workpiece will melt, and finally will be strongly attached on the blade. The diamond tool realized no chip adhesion and no large cutting force, which suggests the importance of sharp cutting ridge.

Absence of tool change and workpiece chuck resulted in long adjusting time. Sometimes we suffered longer adjustment such as 30 minutes. Though the conventional machining process distributes the workpiece to the fixed machines, on the contrary, the micro machining process should distribute the machines to the fixed workpiece. Because fixing or releasing the workpiece / tool is difficult, the micro process prepares many but redundant machines.

Fig. 9 shows a layout of micro machining process. The machines are concentrated around the workpiece ; the workpiece is fixed without releasing during processing : a total environmental control is managed like an umbrella.

6. Summary

To fabricate 1/100 micro parts, a 1/100 machining tool was designed. The machine is composed of many newly developed mechanisms such as a voice coil motor suspended parallel plates spring because some mechanism were very difficult to be miniaturized. The fabricated machine is $26 \times 26 \times 67 \text{ mm}$ in size, 167 g in mass. We could cut the plastic materials with $\phi 50 \mu\text{m}$, carbide endmill fabricated by ourselves using Wire Electro-Discharged Grinding. But we still need a sharp cutting ridge such as a diamond endmill by a $0.1 \mu\text{m}$ edge roundness to cut a plastic material without any burnishing.

We would like to thank THK Co. Ltd. for designing and fabricating the spindle, Prof. T. Masuzawa and Mr. K. Egashira of I.I.S., the Univ. of Tokyo for valuably discussing on the tools made by WEDG.

Reference

- [1] R.P. Feynman, "There's Plenty of Room at the Bottom," J. of Micro Electro Mechanical Systems, vol. 1, 1992, pp.60-66.
- [2] Ed Regis, *Nano, The Emerging Science of Nanotechnology*, Brown and Company, Boston, MA, 1995.
- [3] T. Masuzawa et.al., Wire Electro-Discharge Grinding for Micro-Machining, Annals of CIRP vol. 34, 1985, pp. 431-434.

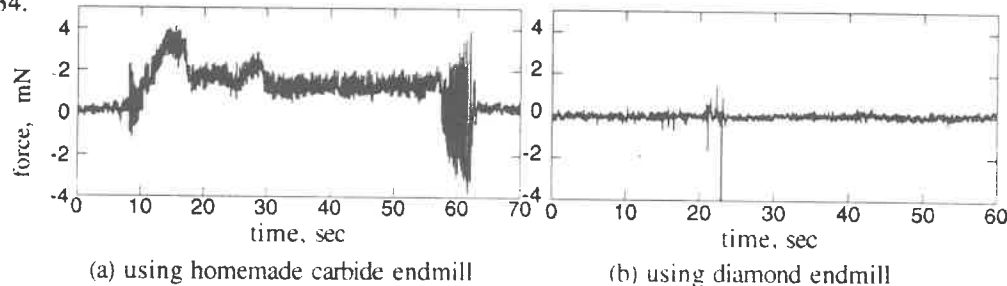


Fig. 8 Measured cutting force in cutting a straight groove
cutting condition : spindle rotation : 60,000 rpm, feed rate : $50 \mu\text{m}/\text{sec}$,
cutting depth : $10 \mu\text{m}$, cutting width: (a) $100 \mu\text{m}$, (b) $380 \mu\text{m}$, workpiece : PMMA

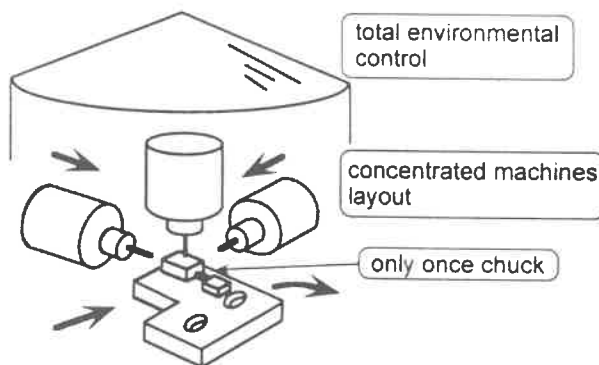


Fig. 9 A new layout of micro machining process

Assembly of micro systems in a large-chamber scanning electron microscope

Manfred Weck, Joachim Hümmeler, Bernd Petersen

Fraunhofer Institute of Production Technology IPT
Aachen (Germany)

INTRODUCTION

More and more micro systems are available which are no longer built up monolithically in silicon, but consist of several different components. These components have to be assembled to form an overall system. Moreover, assembly is required whenever a micro system has to be connected to the outer world or housing is necessary. Today, assembly is usually done by hand with the help of tweezers and optical microscopes. It is obvious that this kind of manufacture is very cost-intensive. Only automation will lead to an efficient and precise assembly process.

PROCESS OBSERVATION IN MICRO ASSEMBLY

In order to better understand what is really happening in the micro world, observation is inevitably necessary. Especially in micro assembly, where precise positioning is essential, the operator must be able to have a close look at the process. Optical microscopy and scanning electron microscopy both offer a "live" view on the production process and are common means of magnification. However, there are some important differences.

An optical microscope is relatively compact and easy to handle. In fact, it is used in a number of applications where micro production is concerned, e.g. the assembly of a micro motor.¹ Compared to a scanning electron microscope (SEM) there are some disadvantages. For example, the lateral resolution and especially the depth of focus are much higher with the SEM. Furthermore, a SEM offers the possibility to work with a large distance between optics and sample (up to several centimeters). This is very important for the observation of micro assembly, as the SEM will leave enough room for a gripper to work below the optics. Another advantage of the SEM is the very clean production environment provided by the vacuum chamber.

The vacuum chamber of an ordinary SEM is very limited in size which in most cases will not allow to study micro production processes. In contrast to this, the large-chamber scanning electron microscope (LC-SEM) "Mira" (VisiTec) offers two exceptional features. First of all, it has a 2 m³ vacuum chamber. This extraordinary large chamber leaves enough space for the installation of machines for micro production. For example at the Fraunhofer IPT cutting, laser welding and assembly have been performed and studied inside the LC-SEM. As a second feature, the electron source can be freely positioned. This allows to look at a production process from all directions while the process itself is left completely unaffected.

MECHANICAL GRIPPERS FOR MICRO ASSEMBLY IN A LC-SEM

In this case the chamber of a LC-SEM serves as environment for the assembly process. Therefore all materials of which the gripper is made have to be suitable for vacuum. For example, they must not release gases, as most plastics do. Another restriction is caused by the electron beam of the LC-SEM, which can easily be deflected by electric or magnetic fields. As this can result in image interferences, these fields have to be avoided. Materials that can be magnetized and electromagnetic drives must not be used in proximity to the electron beam.²

Several different grippers for micro assembly inside a LC-SEM have been developed and tested. Different types of drives, such as bimetallic strips, piezoelectric bimorph elements or shape memory alloys, have been examined. Very good performances were obtained with the gripper shown in figure 1.

The gripper's body consists of a plate made of non-magnetic high quality steel (1.4301). Two metal stripes are attached to this body and form the arms of the gripper. They are 0.2 mm thick and are made of a copper-beryllium alloy (CuBe2). This material was chosen because it can not be magnetized and can be used as a spring (similar to spring steel). A third stripe forms a "U" and is fixed between the arms. With this design no abrasive articulations or lubricated bearings are necessary.³

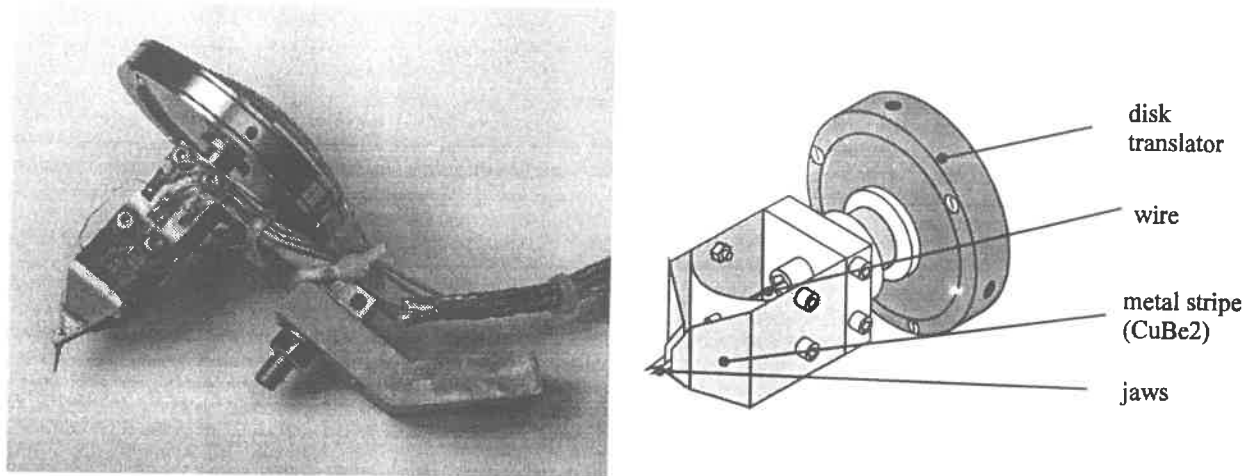


Fig. 1: Mechanical gripper for micro assembly inside a LC-SEM

In order to operate the gripper the "U" stripe has to be deformed by a force, which will cause the jaws to open. After releasing the "U", the gripper will return to its original position with the jaws closed again.⁴ In order to avoid non-symmetrical movements of the jaws - e.g. open jaws but slightly tilted to one side - it is very important that the point of application of force is well in the center of the "U" stripe.

In order to prevent damages to the micro parts caused by a too high contact stress, the control of gripping force is very important. For this reason, resistance strain gauges were fixed to the inner and outer side of each arm. They were attached as near to the jaws as possible, for not being affected by the gripper's deformation due to normal operation. When gripping an element, the arms of the gripper together with the resistance strain gauges will be deformed. Thus a change in electrical resistance is generated which allows to determine the change in gripping force.

A piezoelectric disk translator was used as drive for the gripper. It consists of two thin ceramic disks that are attached to each other. When connected to voltage one of them contracts while the other one expands which causes a bending of the whole disk. The thus generated force is transmitted to the "U" stripe via a thin wire. Piezoelectric elements allow very precise movements and enable the arms of the gripper to move in nanometer steps.

The jaws of the gripper were carefully attached to the end of the arms. For secure gripping it is very important that the jaws come together congruently when the gripper is closed. If this is not the case, a micro part might easily be flicked away when being gripped. Furthermore the jaws have to be very sharp in order to be able to grip elements with the size of only some micrometers. In this case tips of tweezers 50 μm wide and 10 μm thick were used as jaws. The material was titanium, which can not be magnetized.

The above presented gripper is able to move its arms in sub- μm steps and to open its jaws about 2 mm. It reacts almost instantly to any control signal. However, the normal hysteresis of the piezo elements causes a slight hysteresis in movement of the jaws. As for the detection of gripping force, a minimum resolution of 0.5 mN could be achieved.

ASSEMBLY INSIDE THE LC-SEM

An experimental assembly arrangement was set up inside the vacuum chamber of the LC-SEM. The gripper was mounted on a linear motion table in order to let it move up and down (Z-axis). Below the gripper, an assembly platform was arranged on which micro parts could be placed. The platform could be moved in the axes X and Y and rotated around the Z-axis. This arrangement allows to keep the focus of the LC-SEM on the jaws of the gripper all the time. No adjusting of the LC-SEM dependent on movements of the assembly platform is necessary, because with the jaws always visible, all steps of the assembly process can be well observed.

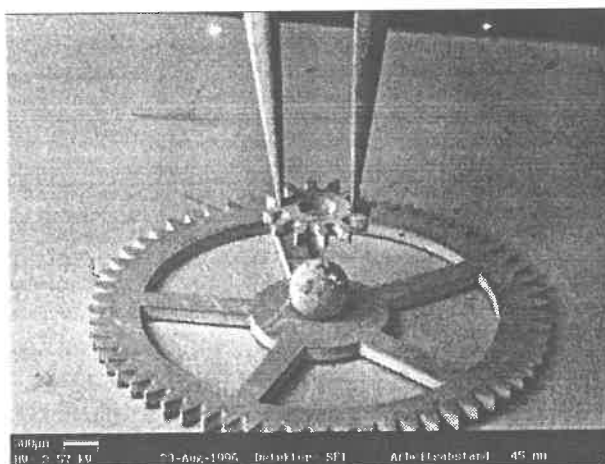


Fig. 2: Micro parts put on top of each other

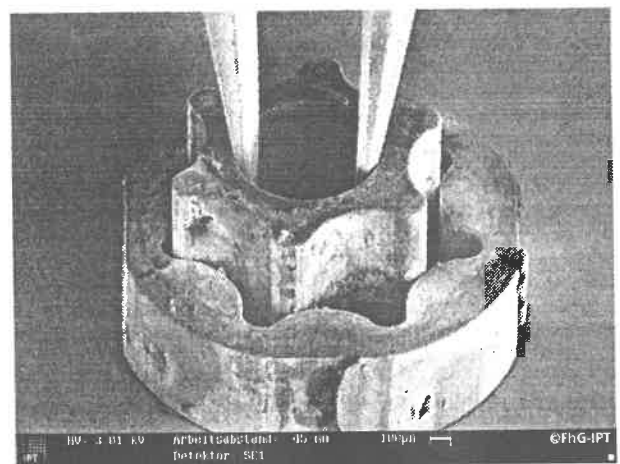


Fig. 3: Assembly of principal components of a micro gear pump

To begin with, some micro parts were put on top of each other in order to test the gripper. Figure 2 shows a gear wheel, a steel ball of 250 μm in diameter and a second gear wheel with a diameter of 500 μm . After that the two principal components of a micro gear pump were assembled (see figure 3). These components were made of hard metal. The outer diameter of the larger part was 2 mm, that of the smaller gear wheel was 1.4 mm. A very high accuracy in

the positioning of the components was necessary because they had to be put together creating a contact in six points with an average gap width of only 1.3 μm .

CONCLUSIONS

It was shown that the large-chamber scanning electron microscope (LC-SEM) is a useful experimental observation system for micro assembly and other micro production processes. It can help examining the phenomena of the micro world. Among other things, components of a micro gear pump were successfully assembled inside the LC-SEM using a remote controlled mechanical gripper. The gripper was designed especially for use in the LC-SEM and was equipped with resistance strain gauges for the control of gripping force. It is powered by piezo elements and can move its jaws in sub- μm steps. With a further integration of sensors in the gripper, the automation of micro assembly will become possible.

ACKNOWLEDGEMENTS

This work was supported by the Collaborative Research Center "Assembly of Hybrid Micro Systems", financed by the Deutsche Forschungsgemeinschaft DFG. The above mentioned micro gear pump was made available by Fraunhofer IPA, Stuttgart (Germany). It was developed for Hydraulik Nord GmbH, Parchim (Germany).

REFERENCES

1. W. Ehrfeld, "Assembly of Electromagnetic Milliactuators with LIGA Components", *2nd Japan-France Congress on Mechatronics*, Takamatsu, Kagawa (Japan), 1994
2. K.-H. Schmidt, "Mikro-Montage im Rasterelektronenmikroskop, dargestellt am Beispiel Nadelsensor", *38. Internationales Wissenschaftliches Kolloquium der TU Ilmenau*, Ilmenau (Germany), 1993
3. J. Dorner, "Braucht die Mikrosystemtechnik die Reinraumtechnik?", *38. Internationales Wissenschaftliches Kolloquium der TU Ilmenau*, Ilmenau (Germany), 1993
4. R. A. Russell, "A robotic system for performing sub-millimeter grasping and manipulation tasks", *Robotic and Autonomous Systems*, vol. 13, 1994

REVIEW OF NON-DESTRUCTIVE MEASURING METHODS FOR THE ASSESSMENT OF SURFACE INTEGRITY

A Survey of New Measuring Methods for Coatings, Layered Structures and Processed Surfaces

G. Goch ^{1,2)}, B. Schmitz ²⁾, J. Geerkens ²⁾, M. Reigl ¹⁾, P. Sprongl ²⁾, R. Ritter ²⁾

¹⁾ Abteilung Meß-, Regel- und Mikrotechnik, University of Ulm, Germany

²⁾ Institut für Lasertechnologien in der Medizin und Meßtechnik, Ulm, Germany

Abstract

Even though the near-surface areas of precisely manufactured parts represent only a few percent of the workpiece's volume, they influence its functional behaviour, quality and lifetime significantly. Therefore, sensors and measuring principles able to detect material changes, damages and process influences in the near-edge zone are of an increasing importance. However, they have to meet industrial purposes concerning accuracy, measuring time, robustness and costs. They should work non-destructively and should offer the suitability for a production process.

As the X-ray, sound propagation and Raman spectroscopy techniques were explained in detail in [1] by Brinksmeier, this contribution takes special emphasis on other non-destructive methods: photothermal, micromagnetic and SNAM-based microhardness measurement. They cover thin near-surface layers from the sub-micrometer to the millimeter-range. Namely the photothermal techniques including various detector-options have gained a substantial progress during the last five years, promising a broad spectrum of measuring abilities. Today, photothermics cover all kinds of layered-structure measurements concerning coatings (thickness, adhesion strength, local disturbances), hardness profiles, residual stresses, wear, thermal influences, and various manufacturing techniques, affecting the near-surface zone. The contribution concludes by the scanning near-field acoustic microscopy (SNAM), allowing simultaneously a surface and a new microindentational inspection.

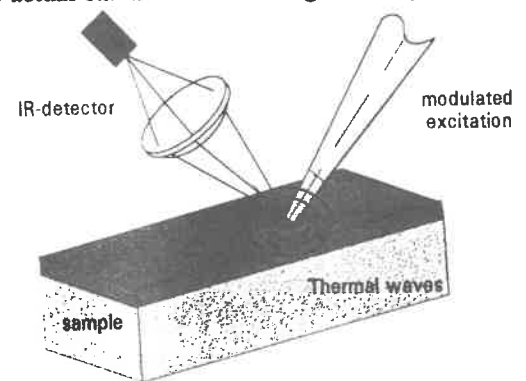
1. Introduction

New capabilities in machining processes, a continuing miniaturization and tolerances down to the nanometer scale have lead to material treatments and depths of cut, affecting only very thin layers next to the surface. The more the over-all volume of a miniaturized product or its affected zone is scaled down in precision engineering, the more these near-surface areas influence its functional behaviour, quality and lifetime, even though they represent only a few percent of the workpiece's volume. Therefore, industrially applicable sensors and measuring principles, detecting material changes, damages and process influences in the near-edge zone, are of an increasing importance. However, they should work non-destructively and meet, at least, the actual standards concerning accuracy, measuring time, robustness and costs.

2. New measuring principles

Since the state-of-the-art of non-destructive measuring methods was summarized by E. Brinksmeier in 1989 [1], special interest is focused on measuring principles suitable for an installation in or near to a production line, i.e. others than pure laboratory instruments. Photothermal, micromagnetic and SNAM-based inspections (see sections 3, 4 and 5, resp.) belong to the indirect methods, correlating the material and structural properties of the near-edge zone with its thermal, electromagnetic or mechanical material parameters and their spatial distribution.

These measuring principles cover thin near-surface layers from the sub- μm - to the mm-range. Additionally, photothermics also include new materials like ceramics and special coatings. All these applications require new measuring and evaluation methods, able to discover new types of deviations (e.g. cracks, adhesion defects, material parameter profiles). Corresponding to the down-scaled thickness of the relevant near-surface layer, the requested measuring resolution and uncertainty must be reduced, too. Due to the importance of the edge zone and a widely stated lack of experiences processing new materials, there is an urgent need to improve manufacturing control, quality inspection and the wear-behaviour concerning all kinds of coatings, edge hardening, sub-surface damages and residual stresses.



UMT-PTR 75.2a

© ILM 1997

Fig.3-1: Excitation of thermal waves and their IR detection

3. Photothermal inspection of the near-surface zone

3.1. Physical principles

In a photothermal experiment, heat is generated by the interaction between an intensity-modulated tightly focused light beam (fig. 3-1) and a sample. Thus, experiments with a resolution in a microscopic scale can be realized. The corresponding thermal response of the periodically heated object can be detected by different sensing devices. The detector signals imply information about the thermal material parameters, layer structures, type and distribution of hidden sub-surface defects. In order to reconstruct the depth and profiles of thermal changes, contrast signals as a function of the modulation frequency f are recorded (frequency sweep). Lateral scans across the sample detect local material deviations.

3.2. Theoretical Description of the Heat Diffusion Process

The basic equation (1) named thermal diffusion equation (TDE) enables to quantify photothermal signals. It describes the heat diffusion process, i.e. the temperature distribution with its spatial and time dependency after a thermal load (exciting laser) [2].

$$\frac{d}{d\underline{x}} \left(\kappa(\underline{x}) \frac{d}{d\underline{x}} T(\underline{x}, t) \right) - \rho(\underline{x}) c(\underline{x}) \frac{d}{dt} T(\underline{x}, t) = -Q(\underline{x}, t) \quad (1)$$

It depends on the spatial distribution of the thermal parameters heat conductivity κ , specific heat c and density ρ . Therefore, general analytical solution can not be achieved, only for special cases regarding some restrictions. For a periodic excitation, thermal waves propagate within the sample. Their penetration depth is given by (2) as the thermal diffusion length μ , which is inverse proportional to the square root of the exciting frequency ω [3]:

$$\mu = \sqrt{\frac{2\alpha}{\omega}} \quad \text{with the thermal diffusivity } \alpha = \kappa / (\rho c) \quad (2)$$

Typical values for μ are shown in table 3-1. Thus, by varying the modulation frequency, photothermal detectors can 'look inside' a surface zone down to a variable depth, without contact, non-destructively, with a microscopic lateral resolution and very quickly.

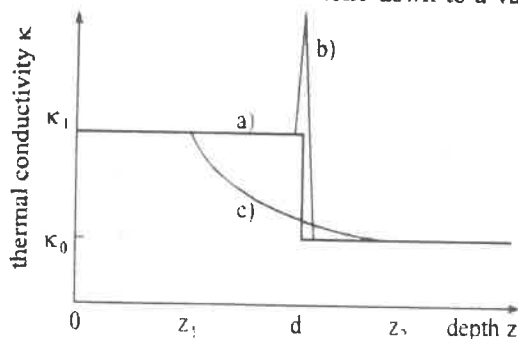


Fig. 3-2: Depth dependence of the thermal conductivity: a) coated system, b) coated system with a thermal contact resistance, c) thermal transition zone [8]

and defect identification belong to the latter application spectrum.

3.3. Layer Thickness and Adhesion Strength

For a non-homogeneous sample, the depth dependence of the thermal parameters (e.g. the thermal conductivity) may follow one of the shapes shown in fig. 3-2 [8]: a simple coating (fig. 3-2a), a thermal contact resistance (fig. 3-2b), modelling e.g. an adhesion defect or a delamination) as well as a continuous transition of the thermal conductivity (fig. 3-2c). This model, combined with a small and robust experimental set-up, has proved already its suitability in many industrial applications. As one example, fig 3-3 shows the calculated (lines) and measured (dots) photothermal phase signals, obtained from a Ni-coated copper sample. Modelling the passivation by a thermal contact resistance, the measurements are in a good agreement with the theory and with a destructive tensile test of the adherence force.

material	diffusion length for 1 Hz	diffusion length for 1 kHz
copper	6.0 mm	190 μ m
steel 16MnCr5	2.0 mm	64 μ m
air	2.6 mm	83 μ m

Table 3-1: Thermal diffusion length for different materials and modulation frequencies

Special problems on adhesion strength occur at the coating of hard-metal or ceramic tool with wear-resistant layers like TiN. Investigations carried out by long-term cutting tests showed that photothermics may offer a regular inspection and prediction of the adhesion quality to be expected for a coating layer, even before this coating was performed.

3.4. Hardness and Residual Stress Measurement

Microstructural changes, as they appear in the steel hardening process occur together with modifications of the thermal properties, i.e. the increasing hardness in a near edge layer causes a decrease of the thermal conductivity in this zone. Therefore, hardness becomes detectable by photothermal means (fig. 3-4).

For the determination of hardness or compound layer thickness, the use of calibration curves (destructively obtained Vickers data versus photothermal data) offers a simple and practical approach.

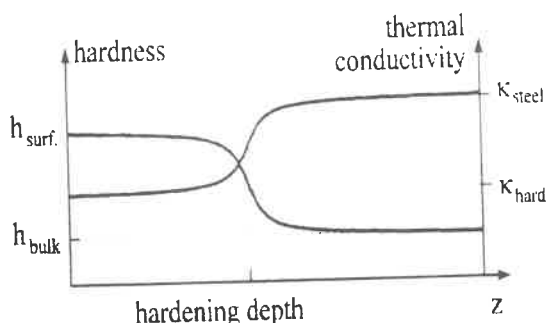


Fig. 3-4: Inverse relationship between hardness and the thermal conductivity

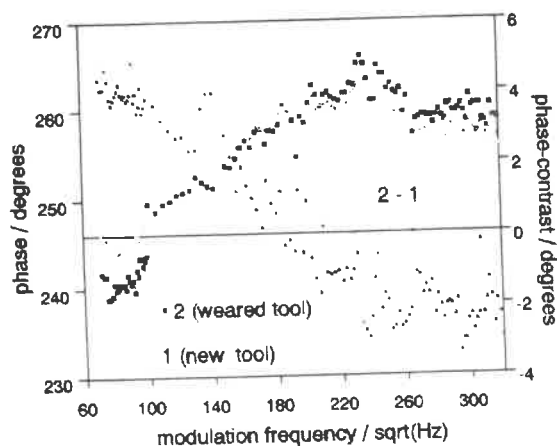


Fig. 3-5: left: Radiometric phase signals of drive shafts with tensile (2) and compressive (1) residual stress, turned with a worn (2) or a new tool (1); right: Phase contrast (difference) signal '2-1'

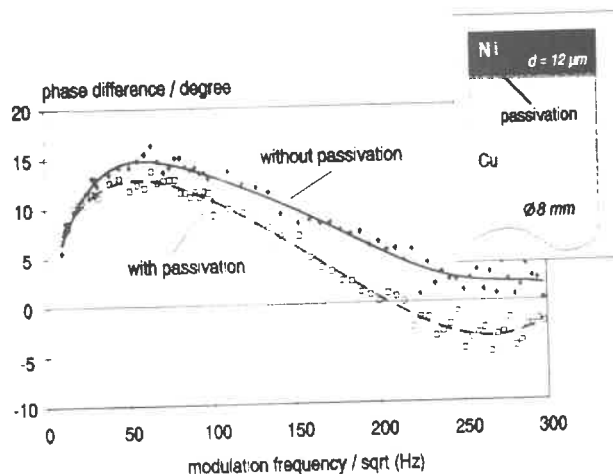


Fig. 3-3: Phase signal of a Ni-Cu sample with a high / low adhesive strength (without / with passivation)

In order to evaluate the shape of the hardness profile, several analytical and numerical models for a theoretical prediction of the photothermal signals were developed in the last five years [7]. The accuracy and application spectrum of these models increase with their mathematical complexity and computation time. Numerical methods like the FDM [6] also offer a 3-D evaluation of the photothermal signals.

Photothermics also enable to detect residual stress in a non-contact and non-destructive manner, in or near to production lines. Fig. 3-5 summarizes radiometric measurements of two hard-turned shafts, one produced with a new turning tool and a second one cut with a worn tool. From X-ray measurements as described in [1], the actual surface stress values and orientation of part 1 (compressive) and 2 (tensile) were known. As a minimum result, the sign of an actual stress condition in a surface layer (i.e. compressive or tensile) can clearly be distinguished by contrast signals, which qualifies photothermal techniques for production control [9].

3.5. Photothermal Investigations of Advanced Ceramics

The final processing of ceramics induce micro-structural changes as well as the micro-cracks influence any heat propagation in the affected region. Therefore, photothermal measuring techniques yield information about the changed mechanical or structural properties of the surface and near surface region. The model-based reconstruction of thermal conductivity profiles from photothermal signals (fig. 3-5) enables to estimate the impacts on mechanically and laser

treated ceramic surfaces. The results justify the assumption that the physical properties of the ceramic samples have been changed in the surroundings of the Vickers indentation (as a 'model' for grinding influences) or laser treated areas [10]. The impacts yield changes of thermal properties and, hence, become detectable by photothermal techniques.

4. Micromagnetic measuring methods

Coercivity, saturation magnetization and magnetic remanence occurring on a hysteresis loop are significantly affected by hardness or mechanical deformation. These well known phenomena can be explained by a lattice distortion and strain, leading to an orientation change of magnetic dipoles, displacements of Bloch walls, formation / junction of Weiss's domains and Barkhausen-jumps. The resulting magnetic Barkhausen noise preferably occurs at the coercive force[10], thus, the recorded noise changes its maximum amplitude and position within the hysteresis loop.

Conventionally, the hysteresis loop is measured with a transformer-like set-up, where the inspected object represents the core. Another, micromagnetic principle records an (amplitude-changed) Barkhausen noise as a function of a (changed) coercivity, reflecting the altered material properties in the near-surface zone. Exciting the sample periodically, the maximum noise amplitudes coincide with the (stress or hardness dependent) coercive force. A depth measurement from the mm- to the μm -range is achievable.

Similar to photothermal measurements, depth profiles of surface layer parameters can be evaluated by frequency sweeps of the exciting and analysing frequency [1]. On the other hand, the restriction to ferrous materials limits the scope of possible applications significantly, as e.g. all ceramics and non-ferrous objects cannot be inspected.

5. Contactless Topography and Microhardness Measurements Using SNAM

Among others, the Scanning Near-Field Acoustic Microscopy (SNAM) can be used as a contactless and non-destructive technique to inspect the topography, profile and roughness of a surface [11,12]. Recently finished projects proved that the same approach is suitable to measure microhardness of thin layers. The main advantage consists in the reduced preparation of the ground, milled or turned surfaces, where polishing is no longer necessary.

The scanning near-field acoustic microscopy belongs to the scanning probe techniques. A vibrating tip attached to a quartz tuning fork works as a distance sensor, damped by the air between the surface and the vibrating tip, which is pressed out and sucked in again periodically. In order to image the surface topography, the sensor scans across the surface in a constant distance (i.e. constant damping), of about 100 nm. Experimental comparisons showed a very good coincidence between mechanical and acoustical tracing.

The SNAM-microhardness approach leads to a three-step inspection procedure. First, the surface topography is imaged at a certain track. Second, the tip is pressed into the surface during its second scan along the same profile. Third, mapping the same line again and subtracting line 1 from line 2 results in the indentation profile alone.

6. References

- [1] Brinksmeier, E.: State of-the-Art of Non-destructive Measurement of Sub-surface Material Properties and Damages, *Prec. Eng.* **11** (4), (1989), 211 - 224
- [2] Carslaw, H.S.; Jaeger, J.C.: *Conduction of Heat in Solids*, Oxford Univ. Press, Oxford, (1959)
- [3] Mandelis, A. (Ed.): *Principles and Perspectives of Photothermal and Photoacoustic Phenomena*, Elsevier, New York (1992)
- [4] Goch, G.; Reigl, M.: Application of the Finite-Element Method and the Finite-Difference Method to Photothermal Inspections, *J. Appl. Phys.* **79** (11), (1996), 9084 - 9089
- [5] Reigl, M.: *Methods for the Quantification of Photothermal Signals*, PhD thesis, University of Ulm, (1997) (in german)
- [6] Seidel, U.; Haupt, K.; Walther, H.G.: Analysis of the detectability of buried inhomogeneities by means of photothermal microscopy, *J. Appl. Phys.* **75** (9), (1994), 4396-4401
- [7] Lan, T.T.N.; Seidel, U.; Walther, H.G.: The determination of microstructural depth profiles by means of photothermal measurements: Theory, *J. Appl. Phys.* **77**, (1995), 4739-4745
- [8] Schmitz, B.; Goch, G.: Charakterisierung von Oberflächen und randnahen Bereichen mit photothermischen Meßverfahren, *Jahrbuch Oberflächentechnik*, **52**, (1995), Ed. A. Zielonka, Hüthig Verlag, Heidelberg-Germany, 239-281
- [9] Goch, G.; Schmitz, B.; Reigl, M.; Reick, M.; Geerkens, J.; Lechleiter, K.: Photothermische Sensorik zur berührungslosen Charakterisierung oberflächennaher Bereiche, *VDI-Ber.* **1255**, (1996) 113-118
- [10] Crostak, H.A.; Bischoff, W.; Polaud B.: Induktive Prüfverfahren, at *Survey 89/1 Angewandte Technik, Wegweiser für die zerstörungsfreie Prüfpraxis*, (1989), 16-25
- [11] Guethner, P.; Fischer, U.CH.; Dransfeld, K.: Scanning Near-Field Acoustic Microscopy, *Appl. Phys. B* **48**, (1989), 89-92
- [12] Goch, G.; Volk, R.: Contactless Surface Measurements with a New Acoustic Sensor, *Annals CIRP* **43** (1), (1994), 487-490

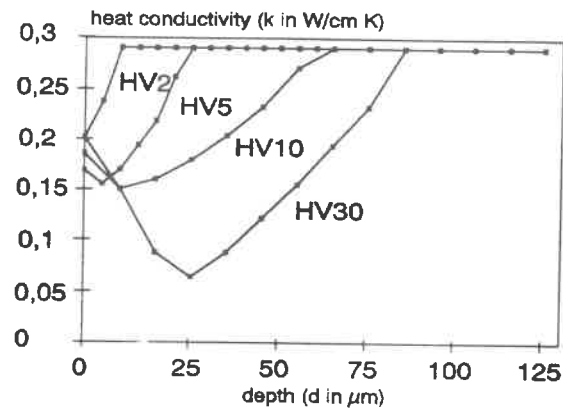


Fig. 3-5: Reconstructed heat conductivity profiles of ceramics, loaded with different forces (Vickers)

Six Axis Active Vibration Isolation and Payload Reaction Force Compensation System

M. Williams¹, P. Subrahmanyam², D. Trumper²

Abstract

"Successful precision engineering is the balance of robustness of the machine and how benign the environment can be made through isolation to minimize the strains caused by vibration that compromises a machine's accuracy." [1]

Photolithography equipment precision positioning capabilities have recently improved to the point where seismic level base disturbances comprise a much larger percentage of the total machine error budget. Additionally, the adverse effects of payload induced reaction forces limit the allowable stage acceleration and thus the wafer throughput of the machine. Current wafer stepper manufacturers have chosen to isolate the system from seismic disturbances and simultaneously profile stage motions to reduce reaction forces. This approach has allowed manufacturers to improve throughput, but is still limited in its capabilities.

We will present an integrated active seismic isolation system with the additional capability of compensating for payload reaction forces, in real time, in six degrees of freedom. The system is comprised of both active and passive elements to allow the isolation or control of disturbance forces in a band which extends from frequencies below 0.5 Hz and beyond any frequency of interest. These two distinct and separate sources of stepper vibration each require a unique controls approach. The vibration induced by stage motions are deterministic in nature and are easily compensated for with feedforward techniques. During wafer exposure, the primary source of vibration is from seismic level base disturbances. These vibrations are usually stochastic and broadband in nature. The fact that the stage motions are deterministic and that the isolation from ground is crucial only during wafer alignment and exposure allow us to hold the system reasonably stationary in inertial space.

In summary, this paper will present the mechanical design and the dual control approach implemented. Ground vibration transmissibility data will also be presented. Additionally, the performance of the system during stage accelerations will be presented. As lithography tools improve in performance, and the industry moves toward very high throughput platforms, these precision isolation and disturbance rejection issues become increasingly more critical.

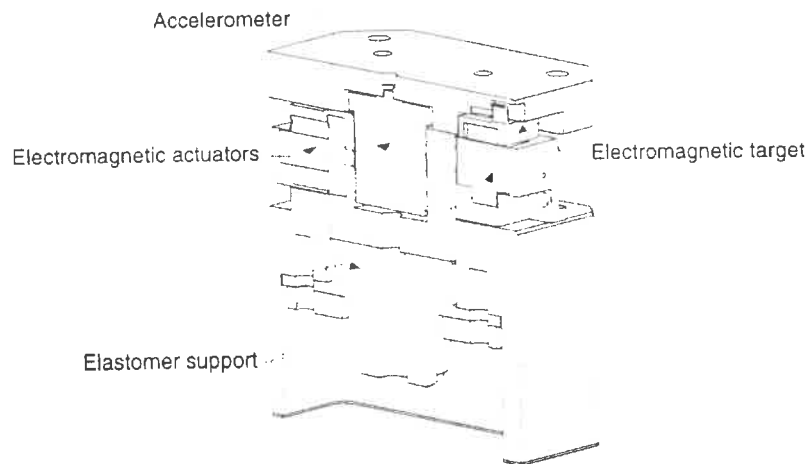


Figure 1: Cutaway view of one leg of the system

¹Integrated Solutions Inc. 836 North Street Bldg 5, Tewksbury, Mass. 01876 williams@mail.insol.com

²Massachusetts Institute of Technology, Dept. of Mechanical Engineering, Cambridge, Mass 02139

Mechanical Design

The isolation system consists of a passive elastomeric mount that supports the mass of the system and also low pass filters seismic disturbance forces above approximately 10 Hz. The shape of the elastomer was chosen to more closely match the horizontal and vertical stiffness. Additionally, damping in the elastomeric mounts is hysteretic, so an increase in damping ratio does not lead to degradation of performance at higher frequencies.

In order to actively control disturbance forces at frequencies below the elastomers capabilities, the system utilizes sixteen variable reluctance tractive type force actuators to control acceleration in six degrees of freedom. These actuators were chosen for their high specific force and relatively large stroke. The system acceleration is sensed and fed back to a digital controller by six seismic accelerometers. These accelerometers were selected based on their high signal to noise ratio even at very low frequencies. Three of the accelerometers have a vertical orientation and sense $\ddot{Z}$, $\ddot{\theta}_x$, and $\ddot{\theta}_y$. The remaining three accelerometers are oriented horizontally and sense $\ddot{X}$, $\ddot{Y}$, and $\ddot{\theta}_z$.

The isolation system actuators are contained in the four supports at each corner of the stepper base. A more desirable arrangement is to use three legs separated by 120° and located closer to the plane that passes through the system center of gravity. This more desirable geometry results in an isolation system that has decoupled rotational and translational natural frequencies. The four support configuration was used to allow a direct substitution for the passive isolation system currently used. Figure 1 is a cutaway view of one of the isolation legs. In this figure, the placement of the elastomer, accelerometer, and actuators can be seen. As seen in Figure 2, the actuators are arranged such that the X and Y force pairs

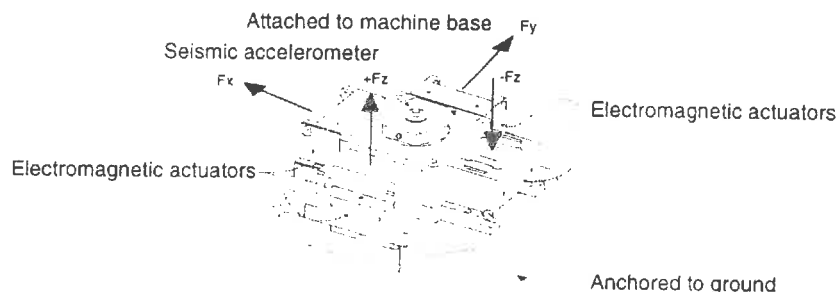


Figure 2: Forces provided at each leg location

are distributed amongst the four legs and the four Z bi-directional pairs are each located on a single leg. The eight pairs of actuators are symmetrically placed and are sized large enough to compensate for a 1-g acceleration of a 30 kg mass.

Modeling and Control Structure

The isolation system was designed based on a model of the stepper as a six degree of freedom rigid body supported by a series of well damped vertical and horizontal springs. The rigid body assumption is valid since the elastic modes of the structure are all well beyond the active control bandwidth. Due to the constraining forces of the static isolation mounts, the system axes of rotation are vertically offset from the centroid and result in tightly coupled equations of motion in $\ddot{X} - \ddot{\theta}_y$ and $\ddot{Y} - \ddot{\theta}_x$. This is not true horizontally since the constraining forces are symmetrical about the centroid. Therefore, the equations of motion for $\ddot{Z}$ and $\ddot{\theta}_z$ are decoupled. The control algorithm is based on this model and is designed to reduce the transmissibility at lower frequencies (a decade below the resonant frequency), suppress the resonant peak at the natural frequency of the mount, and join the passive transmissibility of the mount

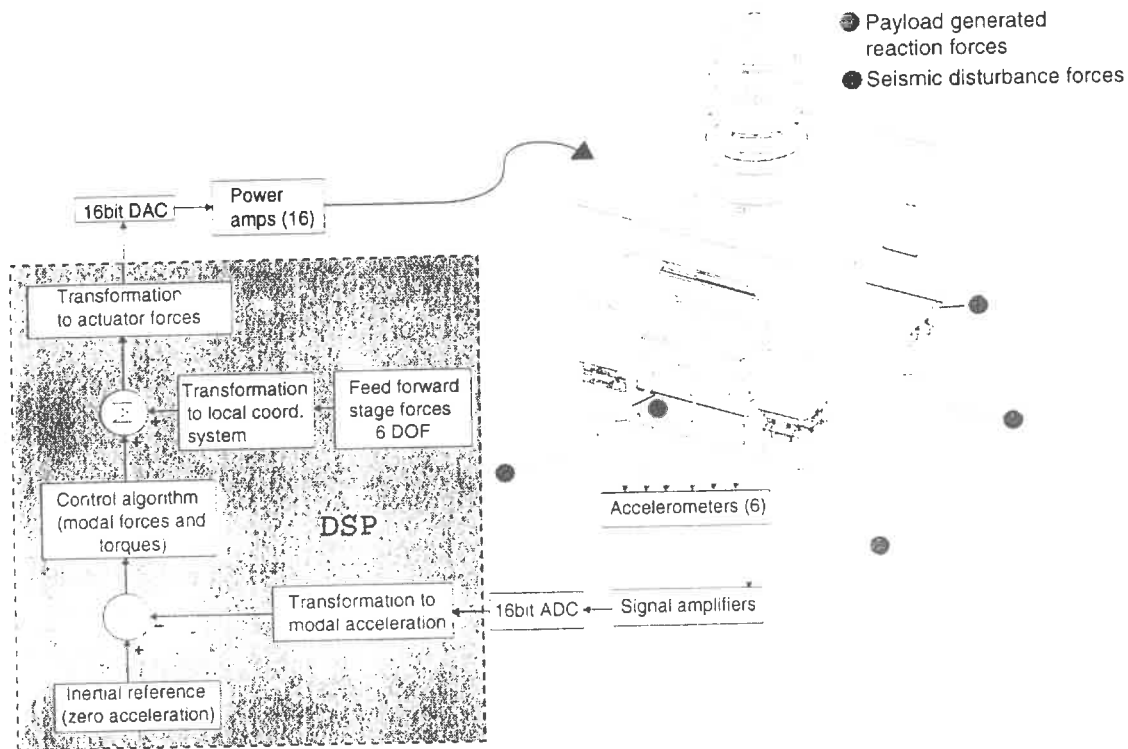


Figure 3: Control structure

about a decade greater than the resonant frequency. Figure 3 is a block diagram representation of the control system hardware and flow.

Results and Conclusions

Preliminary control implementation consisted of two poles. A pole at 1 Hz. to provide 90° phase margin at the low frequency crossover point, and a pole at 70 Hz for 90° phase margin at crossover a decade above resonance. Figure 4 shows the vertical seismic transmissibility of the system with and without active isolation. The figure clearly shows the resonance and the resulting amplification of base vibrations due to the passive mount. Active control is seen to attenuate base vibrations over the control bandwidth. Future work will focus on frequency shaping techniques to provide improved isolation characteristics [2].

In addition to seismic attenuation, the system can also reject payload generated reaction forces. Figure 5 shows the stepper acceleration in the X direction during a 20 mm step of the wafer stage. Without active control, the stepper oscillates at the natural frequency of the passive mount for ≈ 1.5 seconds. With active control the reaction forces are transferred to earth and the stepper is held "quiet" in inertial space.

In conclusion, preliminary results indicate that stage settling times for photolithographic steppers can be improved considerably using the active/passive control system described. Future work will implement MIMO time-optimal control to improve performance by coupling the stage control with the isolation control.

References

- [1] D.B.DeBra, "Vibration Isolation of Precision Machine Tools and Instruments", ASPE short course notes, Cincinnati, Ohio, 1994
- [2] N.K.Gupta, "Frequency-shaped Cost Functionals: Extension of Linear Quadratic Gaussian Design Methods", AIAA Journal of Guidance and Control, Vol. 3, No. 6, 1980. pp 529-535

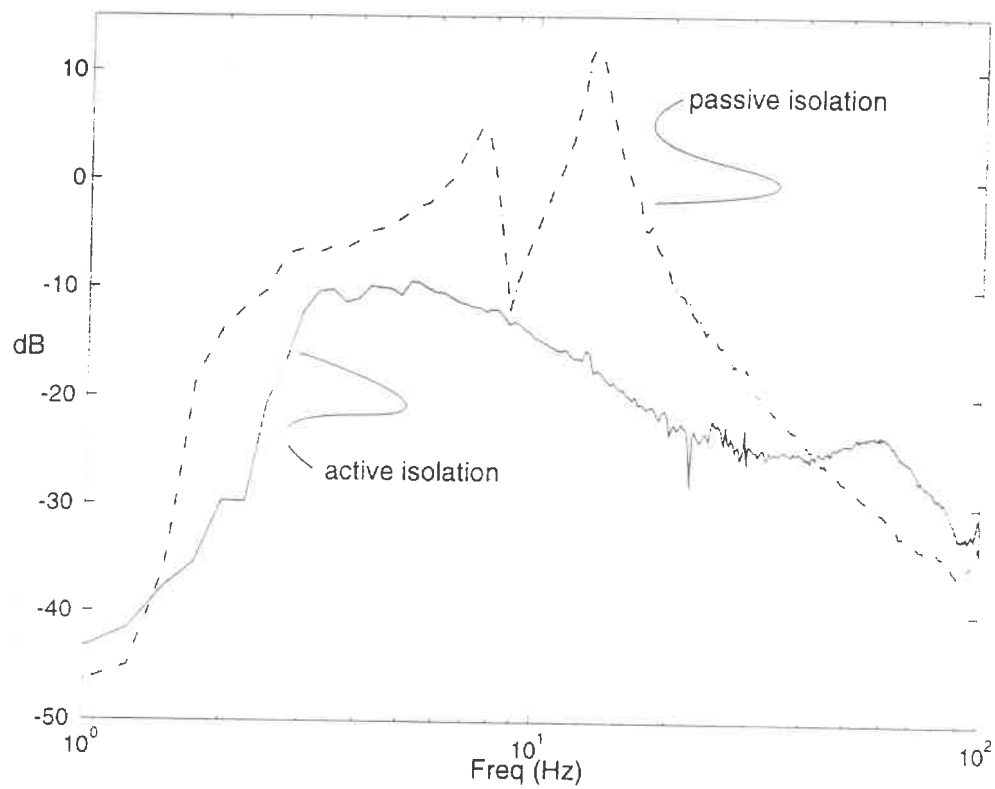


Figure 4: Vertical seismic transmissibility

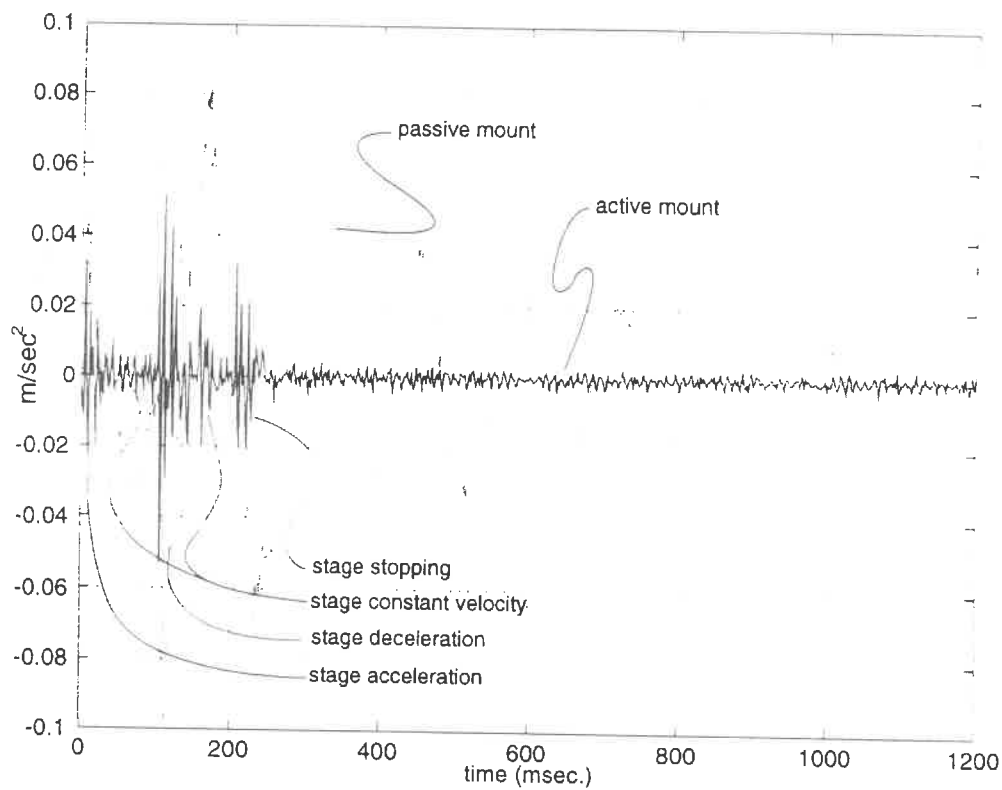


Figure 5: Payload disturbance force rejection

Compensation of Dynamic Vibrations and Spindle-Runout

Manfred Weck, Robert Hilbing

Fraunhofer Institute of Production Technology IPT

Aachen, Germany

Introduction

A primary intention with regard to chip-cutting surface processing of work pieces is to improve the finished surface quality of the work pieces. New approaches are needed to meet the increasingly stringent demands placed on machine tool manufacturers for cost-aware improvement of the machine systems.

Conventionally, machine vibrations are minimized by constructional optimization. This can entail high production costs. With the increased availability of electronic systems and, above all, considering the cost advantage of electronic control compared with mechanical construction optimization, active compensation of relative motion between tool and work piece caused by machine vibration has become an attractive approach to solve the problem. This can be achieved with a position-correcting device, the so-called actuator, on the tool to guide the latter such that the relative distance between the tool and the work piece always corresponds as exactly as possible to the setpoint specifications. Disturbance measurement and calculation of the compensating signal as well as the positioning correction itself require a length of time which is not negligible compared with the period duration of a typical machine vibration. Therefore a predictive projection of the disturbance waveform into the future is required, by the total time taken for measurement, calculation and position correction.

Basic principles

In principle, the possibilities for generating a signal suitable for influencing the vibrational behaviour of machines can be sectioned into control loop algorithms, echo compensation and prediction.

A conventional control loop algorithm is designed to minimize the control deviation with respect to the reference input in a closed loop, by applying a suitable setting variable to the drives and actuators of a machine. A problematic aspect is that the time required for measuring, calculation and position correction is of the order of magnitude of the period duration of the typically encountered, relevant dynamic machine disturbances with vibration frequencies up to 1000 Hz. This means that stable simple closed loop control is not possible.

For echo compensation, a compensatory correction signal is generated for the disturbance variable from the disturbance source signals, in order to achieve compensating position correction during the propagation time of the disturbance origin to the work piece processing location. The disadvantage is here that only complete coverage of all disturbances permits complete compensation.

In the third possible procedure, the correcting signal required to drive the compensator is generated from the disturbance signal itself. In contrast to echo compensation, this no longer requires any knowledge of the actual cause of the disturbance. However, because the principle requires that the compensator correction must be time-coincident with the appearance of the disturbance, but the necessary measurement, calculation and position correction take a finite time, this procedure needs a predictive projection of the disturbance waveform into the future, based on the actual course of the disturbance waveform in the past. Therefore autoregressive process models are used as they are exploited for efficient voice transmission in digital mobile radio communication with the GSM-standard.

In the process model, the human vocal system is represented by a tube of varying cross section to which acoustic pressure is applied by the glottis. The tube is terminated with the lips. Local discretisation of the tube produces the autoregressive digital model of the voice tract. The model of voice generation distinguishes between vowels and consonants. For voiced vowels (A,E,I,O,U) the glottis excites the

voice tract with short pulses, whereas unvoiced consonants (S, SH) are generated by applying noise to the voice tract. For voice encoding, the digital filter parameters and the excitation parameters are determined 20 times per second. The algorithms used for this purpose have been standardized in the digital mobile radiotelephone network.

Application: Vibration compensation in turning machinery

On comparing a consonant such as "SH" with a work piece-tool displacement waveform recorded for a turning machine, it is found that both signals have boosted amplitudes at certain frequencies. These are produced in voice signals by the resonances of the voice tract and for the turning machine by the resonances of the machine structure and the rotation speed dependent characteristic frequencies of the bearings or drive units. The process model assumed for prediction consists of an autoregressive digital filter excited by a noise signal. The filter replicates the resonant behaviour of the voice tract or machine structure. The analysis model which is the inverse of the process model is designated by the term Linear Predictive Coding (LPC) in voice processing systems. Short-term signal prediction by one digital sampling value is accomplished therewith. Multiple step prediction is directly obtainable therefrom by iteration.

Fig. 1 sketches the structural principle of a vibration compensating system for a lathe as set-up at the Fraunhofer IPT. The piezo-electric actuator can be seen interposed between the tool and the tool holder to make the position correction. The disturbance variable (deviation from the proper distance between tool and work piece) is measured with a non-contacting distance sensor and serves as input signal for the mathematical process model. The large measuring surface area of the sensor covering several square millimetres integrates over surface undulations which would otherwise falsify the measurement readings.

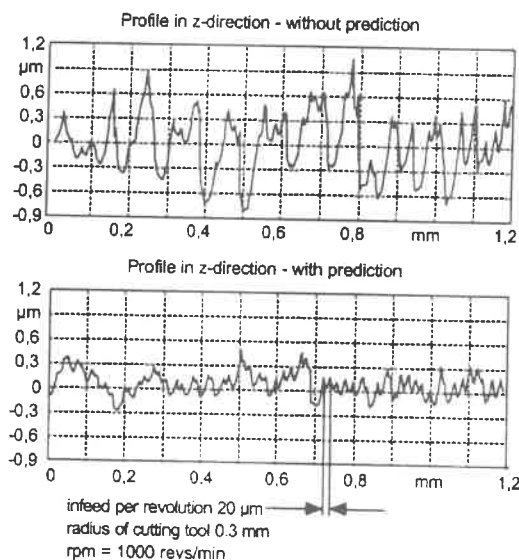
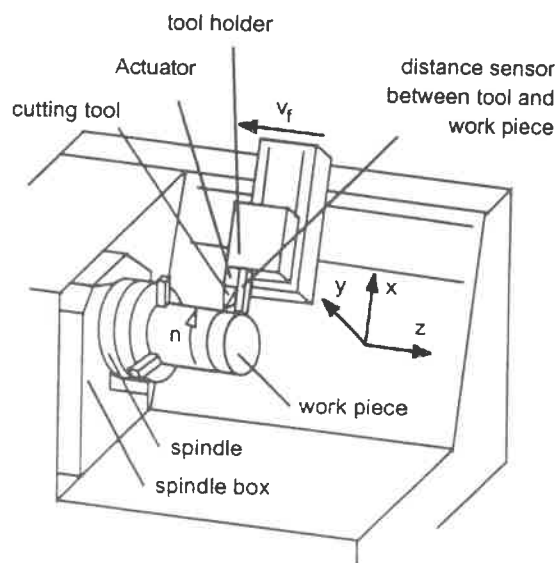


Fig. 1: Sketch showing the structural principle of a vibration compensation for a turning machine implemented in the Fraunhofer IPT.

Fig. 2: Comparison of the surface structure for a turning process with and without prediction

The surface qualities achieved have improved significantly through the use of prediction. For quantisation of the improvement, the surface profile of the work piece was measured with a surface measuring instrument along the z-axis (feed axis). The result is shown in Fig. 2. With the compensation switched off (upper diagram) it is evident that after a few revolutions the tool penetrates deeply in an apparently stochastic manner so that the cutting edge form is clearly impressed. In contrast thereto, nearly all feed grooves are visible in the lower diagram. The amplitude is considerably reduced there, corresponding to a roughness reduction from $R_a = 290 \text{ nm}$ to $R_a = 120 \text{ nm}$.

Application: Eccentric spindle rotation errors.

Eccentric spindle rotation errors are another important factor affecting the quality of the processed work pieces. They arise, for example, through inaccuracies in the spindle bearings and have the effect that the work piece is not truly round. This disturbance usually depends on the current rotation angle of the spindle and is non-statistical to a high degree. The associated dependence on the rotation speed has the consequence that the frequencies of this disturbance, e.g. 33 Hz for a rotation frequency of 1000 revolutions per minute, lie well below the frequencies of the dynamic disturbances considered so far.

A sensor mounted directly onto the tool to measure the relative movement between work piece and tool cannot be used to measure the deviations from roundness. Instead, a reference work piece is required which is rigidly attached to the spindle and has a high quality surface. Such a reference work piece may only have maximum form deviations considerably smaller than the desired form accuracy of the work piece which is to be processed. With the aid of a capacitive sensor, the relative movement of the reference surface in relation to the spindle box is measured in the sensitive direction. The deviation from ideal roundness found in this manner can be used to generate a compensation signal for active position correction.

The first practical implementations of this procedure at the Fraunhofer IPT confirm the fundamental feasibility of compensating deviations from roundness (Fig. 3). The very small round running error of $0.75\text{ }\mu\text{m}$ of a turning machine was further reduced by the described procedure to values around $0.45\text{ }\mu\text{m}$.

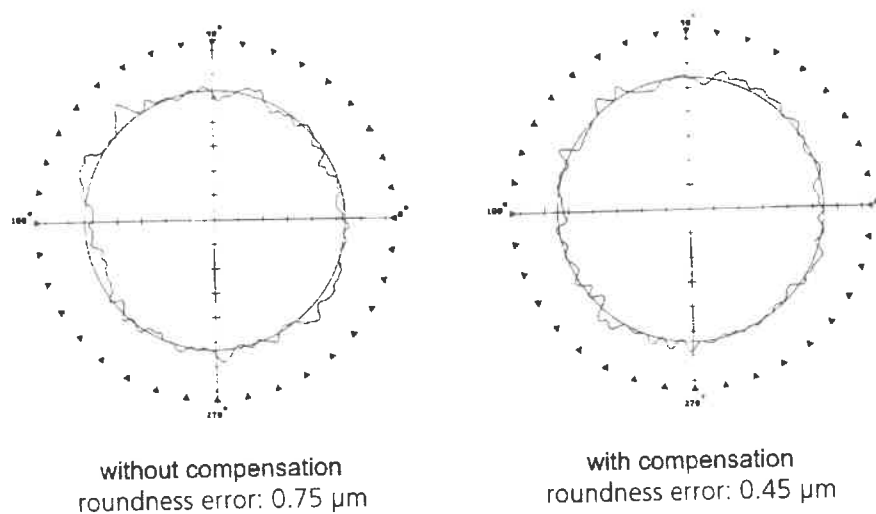


Fig. 3: Reduction of the roundness error by active position correction of the tool (scale subdivision: $0.50\text{ }\mu\text{m}$ / graduation)

Optimization: Using a force sensor

Capacitive sensors are only conditionally suitable for production process measurements because these sensors respond sensitively to cooling lubricants and ejected chip and because the measurements cannot be made directly at the working location, as the sensor has to be mounted adjacent to the tool. An alternative measuring method is to determine the processing forces arising in the position correcting direction. With the cutting parameters involved here, this passive force is approximately proportional to the actual chip thickness. Therefore by measuring the force it is possible to obtain the information about relative deviation of tool and work piece separation required for prediction.

A piezo-electric element is used as sensor for this purpose. It is attached in the force direction and generates an electric charge signal proportional to the applied force. Such a sensor is capable of measuring forces ranging from well below one Newton up to several kilo-Newton with high resolution and ade-

quate bandwidth. It is mounted together with the piezo-electric actuator in the same casing which is completely closed with respect to the exterior.

Thereby the aim is to keep the passive force as nearly constant as possible, because this will achieve uniform chip thickness. This is possible with a modification of the employed algorithms. One part of fig. 4 shows the frequency spectrum of the passive force in the course of normal processing. The typical boosted amplitudes at certain frequencies as characteristic for processing with continuous chip cutting are clearly recognisable. The spectrum becomes significantly flatter after switching-on the compensation algorithm; the disturbance frequency components are suppressed. The surface improvements achieved therewith correspond to a reduction of the roughness R_a by about 30 % compared with processing without compensation.

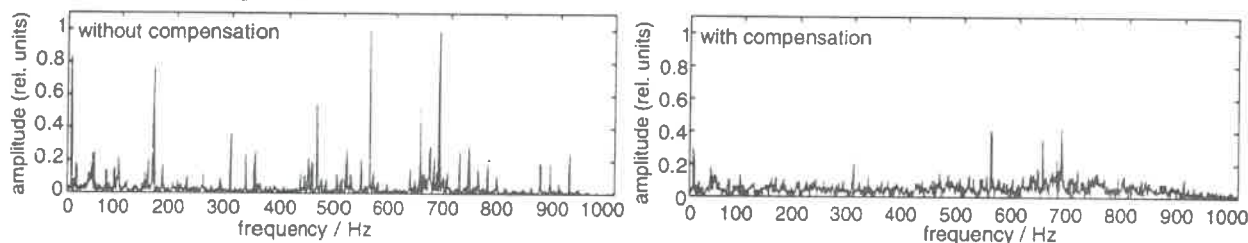


Fig. 4: Comparison of the frequency spectra for processing with and without force compensation

Conclusions and further prospects

In summary it can be said that new position correcting principles such as piezo-electric actuators provide the possibility of compensating high frequency vibrations typically encountered with machine tools, by active tool position correction. However, the propagation delay times of the measuring and positioning path require knowledge of the deviations which would appear without compensation earlier than their actual appearance. Predictive methods are already commonplace in other disciplines such as voice signal processing. It is possible to project vibrations into the future using echo compensation approaches or prediction methods. The investigations carried out for predicting and then compensating the deviations confirm the viability of such methods. By utilising special sensor systems, it is possible to include surface improvements in the production processes.

Apart from the chip-cutting processing machines considered here, special significance is attached to developing further applications in other fields. As far as the machine is concerned, industrial usability can be ensured in future by fixed integration of the sensor systems in the machine and actuators adapted to the production process. It is to be expected that the prediction accuracy can be improved further, by no longer approximating non-linear deviations with linear models, using non-linear prediction models and calculation algorithms instead. Special importance continues to be attached to enhancing the rugged stability of the compensation chain.

This work has been supported by the Volkswagen Foundation.

References

- 1 Kammeyer, K. D., Kroschel, K.: Digitale Signalverarbeitung, Teubner Studienbücher, 1989
- 2 Weck, M.: Werkzeugmaschinen, Fertigungssysteme Bd. 4, Meßtechnische Untersuchung und Beurteilung, 5th edition, VDI-Verlag, Düsseldorf 1996
- 3 Haykin, S.: Adaptive Filter Theory, 3rd edition, Prentice Hall, Upper Saddle River, 1996

1. Combination of mechanical springs with positive and negative stiffness. Disadvantages are the small assembly tolerances and the limited bandwidth due to the resonant frequency of the springs (usually between 10 Hz and 100 Hz for leaf springs).
2. Counterweight balancing. This method doesn't meet the needs of most bearing systems and does not give vibration isolation.
3. Lorentz actuator (also known as 'voice coil actuator'), where $F = B \cdot \ell \cdot I$ so $\partial F / \partial z = 0$, provided that $\partial B / \partial z = 0$. However, for larger bearing forces, energy dissipation and thus thermal problems become substantial.
4. Permanent magnet bearing unit. This new construction element will be described in this paper.

1.3 Possible applications of a zero stiffness bearing unit

In general, three categories of applications can be discerned:

1. Gravity compensation. An arbitrary bearing force can be applied to the object to compensate for $F = mg$ without adding stiffness or damping.
2. Vibration isolation. The new bearing unit is inherently insensitive to the relative distance between the supporting frame and the supported object (within a certain range).
3. Bearing and guidance systems that have a high demand on accuracy. With the new zero-stiffness bearing unit a very stable fine-positioning stage can be built on top of a coarse-positioning system.

The last item includes applications in which an object (e.g. a manipulator table) has to be positioned relative to a target (e.g. a lens), where no disturbing force is allowed on the target, e.g. in microscopes or IC lithography machines.

2 Working principle of the permanent magnet bearing unit

The working principle of the zero-stiffness permanent magnet bearing unit (ZSMB) is based on combining positive and negative stiffness of mutually repulsive and attracting permanent magnets, as depicted in

figure 2 and 3. Thanks to modern magnet materials (e.g. rare earth metals) it is possible to generate a strong repulsive force without demagnetizing of the magnets. When properly designed, it is possible to create a working point where the stiffness in the bearing direction is virtually zero. Within some region, the stiffness can be very low, as is shown in figure 4. The lateral stiffness in the working point is negative, but can be kept small by carefully designing the magnet geometry.

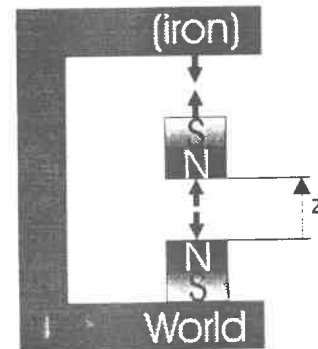


Figure 2: Principle of the zero-stiffness bearing unit using permanent magnets. The attractive force (above) and repulsive force (below) create negative and positive stiffness, respectively and cause the total stiffness to be zero in a working point.

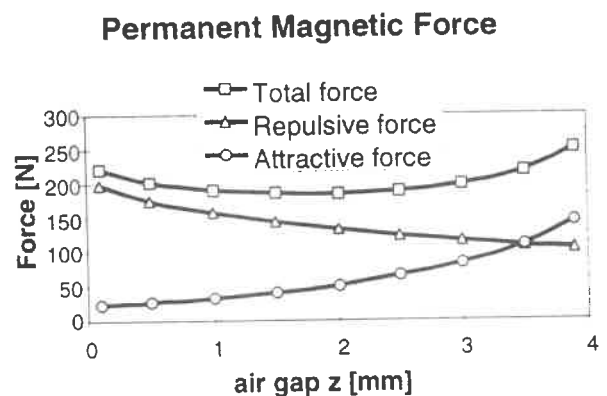


Figure 3: When the attractive and repulsive forces of the permanent magnets are added up, the total force curve shows a minimum where the first derivative – the stiffness – is zero. The values in this figure are obtained from a Finite Element Analysis using the model shown in figure 6.

The resulting bearing force is adjustable:

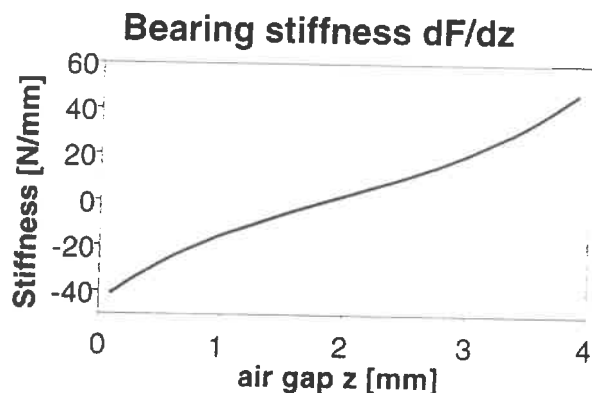


Figure 4: The curve representing the first derivative of the total force (i.e. the stiffness), which is zero in a working point and near-zero in a small region.

1. by changing the air gap (for static adjustments), or
2. by an active control system, for example by adding a coil to generate an additional force on the supported object.

The latter adjustment can be applied to compensate not only for low-frequency drift, but also for higher-frequency disturbing forces that might act on the supported object from the outside. This way, it is possible to create a high (virtual) stiffness bearing without any mechanical coupling to the supporting base.

2.1 Stability of the bearing force

As can be seen in figure 3, when adding the attractive and repulsive force curve, there is a minimum in the total force. That is also the desired working point, because at that point $\partial F/\partial z = 0$. The neutral equilibrium in this point can be shown to be marginally stable. This can be explained as follows:

Suppose the weight of the supported object is slightly more than the minimum bearing force. Then there exist two intersections of the line $F = mg$ with the bearing force curve, of which one (A) is stable and the other (B) is unstable, see figure 5. This can be represented graphically by a rolling ball on a curved surface under the influence of gravity. The additional active controller can be considered as a rotating mechanism providing neutral equilibrium for the rolling ball, that is, $A=B$.

Summarizing, zero stiffness bearings enable supporting an object (around a certain working point)

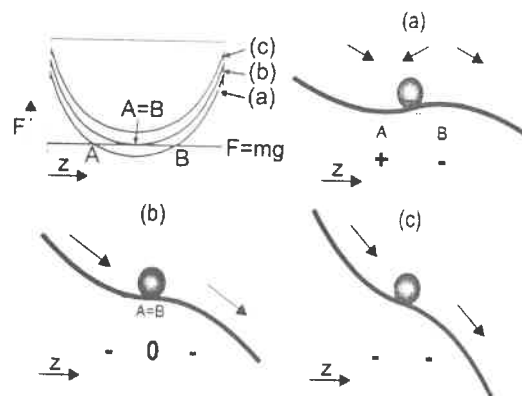


Figure 5: Analogy for the stability in the working point. The horizontal ball movement corresponds to the displacement z in figures 2 through 4. (a) bearing force is smaller than weight $\rightarrow$ stable in A, unstable in B; (b) bearing force equals weight $\rightarrow$ neutral equilibrium only in the working point; (c) bearing force is larger than weight $\rightarrow$ unstable.

in such a way that:

- gravity is compensated for by the position-independent bearing force
- only inertia effects act on the object
- vibrations from the outer world with limited amplitude do not influence the object's position (except for acoustic coupling)
- arbitrary virtual stiffness relative to any other object is possible by means of active control, yet independent of the supporting construction

2.2 The ZSMB compared to the Lorentz actuator

Comparing the zero-stiffness permanent magnet bearing unit to the well-known Lorentz actuator or voice coil (generating a force proportional to the input current), the following qualifications can be formulated (+ means favorable):

	Lorentz act.	ZSMB
Maximum force	-	++
Force bandwidth	++	-
Isolation bandwidth	++	++
Energy dissipation	--	++
Stiffness = 0	++	+

The ZSMB appears to be a good replacement for Lorentz actuators in low-frequency (bearing) applications. In other words, the ZSMB can be regarded as a DC module for a position-independent force, without energy dissipation. This position-independency makes it very suitable for modular use as well, i.e. superposition of the force with another force is possible without mechanical interaction.

3 Experimental setup

According to the principle of figure 2, a practical setup has been designed as shown in figure 6. The dimensions of the applied magnets are $\varnothing 25 \text{ mm} \times 10 \text{ mm}$. The behavior was simulated using a Finite Element Analysis, resulting in the flux lines included in figure 6 and the force curves as shown in figure 3.

The slope of the stiffness curve around the working point is about 11 N/mm^2 , so for a working region of $\pm 5 \text{ } \mu\text{m}$ the maximum stiffness will be 55 N/m . That means that a mass of 18.9 kg (corresponding to the bearing force of 185 N) has a first resonance frequency lower than 0.27 Hz (for comparison: to achieve similar results with a spring, the minimal elongation should be $\frac{185 \text{ N}}{55 \text{ N/m}} = 3.4 \text{ m!}$).

Figure 7 gives a possible test setup to measure the relation between force and displacement experimentally. While the force sensor, attached to the middle magnet, is moved along the desired route, the force and the displacement are measured simultaneously. In this way the deflection of the sensor (e.g. a strain gauge type sensor) doesn't influence the measurement.

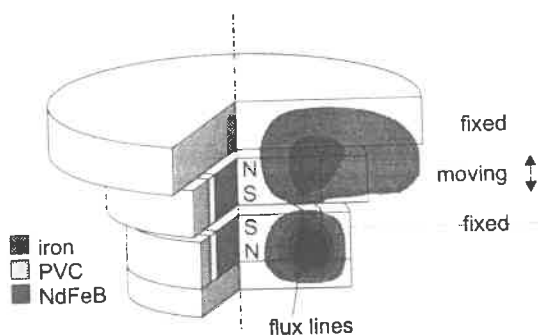


Figure 6: Experimental setup, consisting of two permanent magnets with some flux guidance. The middle magnet is moving, the upper and lower part are fixed to the world. In the right half of the figure the flux lines are shown, as obtained from a FEM analysis.

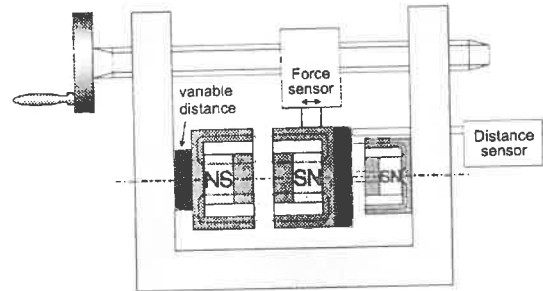


Figure 7: Schematic test setup. The middle magnet is moved from left to right, while the displacement and the force are measured simultaneously, yielding the curve $F = f(x)$.

4 Design of a 6-DoF ZSMB

For six-degrees-of-freedom the approach as sketched in figure 2 will not suffice, because of the lateral non-zero stiffness. Therefore, a more general method has to be used, analytically deriving the required bearing properties from the six-DoF zero-stiffness conditions.

It appears that, under certain conditions, it is indeed possible to realize a bearing system based on permanent magnets giving zero stiffness in all six degrees of freedom. It is obvious that active control will be necessary to prevent drift of the supported object, but this doesn't require high bandwidth.

5 Conclusions

1. With the help of permanent magnets it is possible to create a considerable bearing force, while the bearing stiffness is zero in a working point.
2. The zero stiffness permanent magnetic bearing is very suitable for vibration isolation and for applications in which the force coupling and the displacement coupling have to be separated.

6 Note

A Dutch patent application is in preparation.

Mechatronics Analysis and Design of High-Precision Drives

E. Kallenbach⁺, P. Maißer⁺⁺, J. Lückel⁺⁺⁺, E. Saffert⁺, C. Schäffel⁺, H. Freudenberg⁺⁺, W. Kuhlbusch⁺⁺⁺

⁺ Technical University of Ilmenau, Faculty of Mechanical Engineering, Institute of Microsystems Technology, Mechatronics and Mechanics, P.O. Box 100565, D-98684 Ilmenau, Germany

⁺⁺ Institute of Mechatronics at Technical University of Chemnitz-Zwickau, Reichenhainer Strasse 88, D-09126 Chemnitz, Germany

⁺⁺⁺ University-GH Paderborn, Department 10 - Mechanical Engineering, Mechatronics Laboratory Paderborn Pohlweg 55, D-33098 Paderborn, Germany

Abstract

The paper describes the joint effort of three teams within the project Modeling, Simulation and Validation of Integrated Electromechanical Multi-Coordinate Drives. The project funded by the German Research Association lays the foundation for analysis and design tools that will assist engineers in the design of high-speed high-precision multi-coordinate drives.

1 Introduction

Increasing precision and faster processes are crucial factors in modern high technology. This area of technology requires, therefore, new drive concepts for motion in multiple coordinates with a positioning accuracy in the submicrometer range and high speed [1]. Applications include wafer lithography, measuring and inspection based on image processing, circuit testing, bonder heads, laser processing and many more.

2 Drive Structure

Figure 1 gives an overview of the drive design. An aero-statically supported stage carries symmetrically arranged magnetic circuits consisting each of two NdFeB-magnets and an iron yoke being closed by the iron stator. The moving magnet concept guarantees a high stiffness of the air-bearings which move on the stator that is manufactured with a high precision flat surface. The flat motor coils are again arranged symmetrically for the x- and y-axes. Opposite pairs of coils generate the force in the coordinate that is

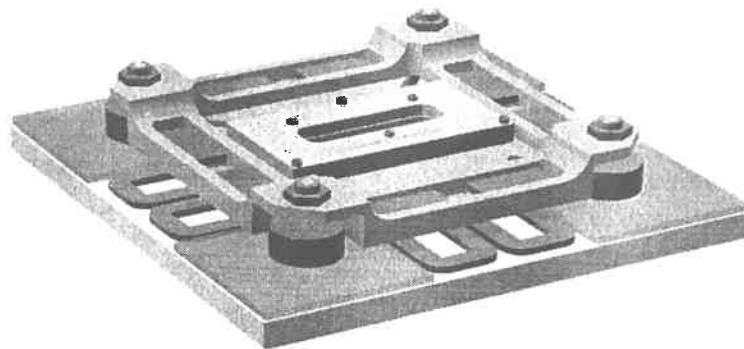


Figure 1. Mechanical structure of multi-coordinate drive

perpendicular to the coil arrangement. Different currents in opposite coils will generate a torque and allow thus to control the stage's angular position around the z-axis. The stage also carries a cross grid. It is scanned by an integrated measuring system for x, y and ϕ which is located in the center of the stator beneath the cross grid. A detailed description of the design including the force generation and the measuring system can be found in [1,4].

3 Modeling and Simulation of Mechatronic Systems

The simulation tool **alaska**, developed at the Institute of Mechatronics Chemnitz, Germany can be used for such problems [2]. **alaska** is based on a unified mathematical description of discrete electromechanical systems by using Lagrange's approach for multibody and electrical systems. A multibody system is a finite set of rigid/flexible bodies embedded in the three-dimensional Euclidean space. The bodies are coupled physically and/or geometrically to each other by force laws and constraints, respectively. The application of Lagrange's approach to electrical systems with lumped parameters is based on the concept of a multipole and its representation by abstract 2-poles as well as Kirchhoff's theory. Methods of graph theory are used to characterize topological properties of electrical networks. An electromechanical system can be understood as a finite set of physical objects with mechanical and/or electrical multipole-properties interacting among themselves by electrodynamic and/or electro-magneto-mechanical coupling.

The coupling of the mechanical and the electrical substructure will be described by constitutive equations containing mechanical and electrical generalized coordinates simultaneously. **alaska** automatically generates the equations of motion for the discrete electromechanical system. All of the electromechanical interactions are correctly taken into account. Even measurement systems, controllers and hydraulic/pneumatic components can be included in the modeling and simulation done with **alaska**. The (differential-) equations describing the dynamic behavior of such components are combined with the equations of motion of the electromechanical system. Sometimes the model-description language is not sufficient for the description of complex technical processes. In this case, **alaska** has an interface for using external FORTRAN-routines. The time-discrete behavior of controllers can also be described by such routines which is important for the simulation of real processes.

The dynamic behavior of a mechatronic system is influenced by its system components. If all system components and interactions can be described uniformly using only one simulation tool, then the dynamic behavior of the whole system and the influence of several components of the whole system can be analyzed. The main part of the mechanical substructure of the multi-coordinate drive is the stage. A rigid and a flexible model of the stage were generated. The flexible model is based on a complex FEM structure reduced to a structure with a few master nodes only. The reduced structure has the same elastic properties as the complex structure. A spatial force coupling element can be defined in **alaska** using the stiffness matrix of the reduced model. The multibody system model of the flexible stage contains 8 rigid bodies connected by this spatial force coupling element. A comparison of the multibody system model and the complex and reduced FEM model with respect to the eigenfrequencies gives a good agreement. The main part of the description of the electrical substructure in this special case is the calculation of the magnetic flux in the coils. This requires the calculation of the normal and tangential components of the magnetic induction between permanent magnets and yoke for any position of the stage. If boundary effects are neglected Maxwell's equations $\text{div } \mathbf{B} = 0$ and $\text{rot } \mathbf{B} = 0$ give a finite Fourier series as solution. The graph of the electrical network consists of 8 independent loops each containing a current generator and an inductance and a resistor. All current generators can be controlled separately. For a continuous stage movement the coils are supplied by commutated currents using suitable time dependent functions. The model of the measurement system contains the periodic sampling, filtering and transfer of measured values and the influence of disturbances. This model is implemented via a special routine. For controlled stage motion, state controllers with an integrated observer are implemented in the model. The modular structure of the whole model enables an easy exchange of submodels, for example the implementation of other controllers.

Simulations of the whole model allows to judge the dynamic behavior of the system and to investigate the influence of system components on it. The range of dimensions of the multi-coordinate drive is very different, the air gap of the bearings is in the range of μm , geometric dimensions of the stage are in the range of centimeters. This requires a high accuracy in the numerical analysis with a long integration time. In view of effective computation and using these models for a virtual prototyping a short integration time is of interest. For this aim the models of system components will be tuned, i.e. looking for simplified models with the dynamic behavior of complex models (simulation results with same exactness). In the near future symbolic codes of the models can be used for nonlinear optimization (optimal layout) and also with respect to the synthesis of new controllers.

4 Path Control

The aim of the control concept is to achieve high accuracy as well as good vibration damping for path oriented movements at high speed. To meet these demands possible disturbances must be considered in the control scheme. Disturbances may occur due to external forces or due to unmodeled system behavior of the plant.

The controller design consists of a structure optimization followed by an optimization of controller parameters. For the synthesis of the controller structure a symbolic model of the plant is used. The appropriate modeling depth of this symbolic model depends on the control accuracy required, but due to handling constraints it cannot be of any complexity.

4.1 Symbolic Design of the Controller Structure

The design starts at the main structure (figure 2) which is defined by a required accuracy of the controlled system as well as by the arrangement of actuators and sensors. Each controlled axis (x, y, ϕ) consists of a local state controller and a disturbance estimator. Furthermore a feed-forward control is included based not only on position but also on velocity and acceleration which are also provided by the path generator. The mathematical representations of some controller blocks are determined in a structure optimization process making use of a procedure based on the symbol-

manipulation program MAPLE. This procedure allows the symbolic coupling and analysis of linear state-space systems. The aim is to obtain feed-forward gains and filters for the observed disturbance which are nearly independent of the plant. The analysis of the reference and the disturbance transfer functions of the coupled system with an iteration process leads to an optimal controller structure. During the iteration the models of the disturbance estimator and the disturbance gains are varied. In this case the result is that the feed-forward gains are independent of any plant parameters if an estimator is used whose disturbance model consists of one or two integrators.

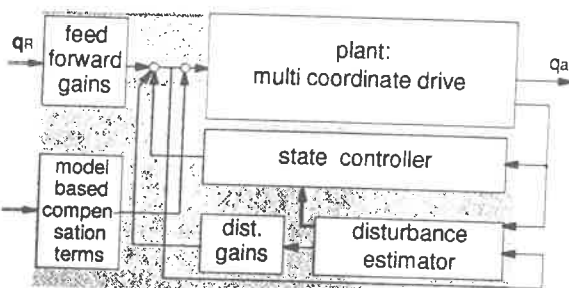


Figure 2. Main structure of the controller

Attention is also directed to the compensation of disturbance forces due to a non-symmetric stage mass distribution because of a working tool mounted on the stage. Since there is no mechanical anti-rotation lock the moment generated by the stage's acceleration will cause a rotation of the stage which is subsequently compensated by the ϕ -controller. This undesirable behavior can be prevented or at least considerably decreased by inserting a model based feed-forward control. This assumes that the parameters of the tool (mass, inertia, distance to the center of gravity of the stage) are well known and can be considered in the control scheme. If the tool parameters are unknown the disturbance estimators are used to adaptively determine them under the assumption that other disturbances can be neglected. In this case the inertia cannot be determined so that the tool's mass must be assumed as a concentrated mass. The detected forces are fed forward via the block *model based compensation terms* as shown in figure 1. The compensation of the nonlinear force-position-characteristic of the drive is done mainly by means of the inverse of a simplified trigonometric series which approximates the magnetic flux through the coils. The error caused by the simplification is detected and compensated by means of the disturbance estimator.

4.2 Optimization of the controller parameters

Once the controller structure is determined its parameters have to be optimized. For this the CAMEL tool [5] is used which was developed at the Mechatronics Laboratory Paderborn (MLaP). The free parameters are optimized using MOPO (multi objective parameter optimization [6]). This leads to a *pareto-optimal* solution minimizing a vectorial quality criterion. The objectives are calculated via nonlinear simulation of the closed loop system which is excited by a suitable step function. By this any nonlinearity is considered including the limit of the servo unit. Besides the controlled quantity the tilt angles are used as objectives as well as the horizontal velocities to improve damping. During the optimization procedure the use of frequency and damping parameters turned out to be more efficient than the use of the feedback gains themselves. Moreover these controller parameters are good starting points for optimization of similar plants with other parameters.

4.3 Simulation Results

The effectiveness of the designed controller is underlined by the simulation results. Both the x- and the y-axis are excited by a path generator

with the following maximal values: jerk 300 m/s^3 , acceleration 6 m/s^2 , and velocity 0.25 m/s . The movement follows a rounded square 30 by 30 mm. At any time the position error of both axis is less than $0.2 \mu\text{m}$ and even less than $0.02 \mu\text{m}$ if the acceleration torque caused by unsymmetric mass

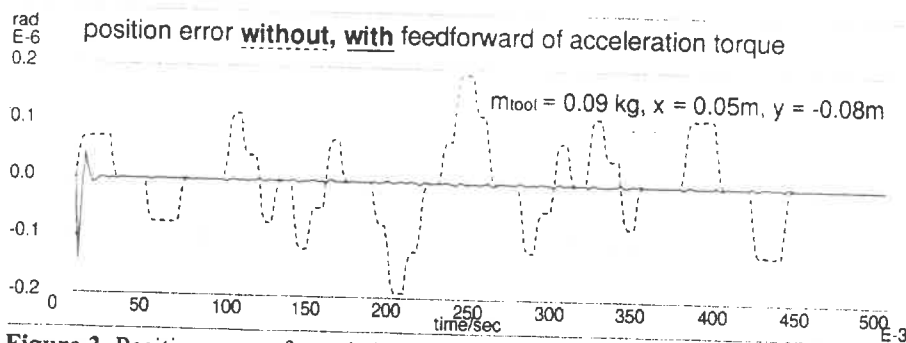


Figure 3. Position error of ϕ -axis in case of unsymmetric mass distribution

distribution is fed-forward. Figure 3 shows the corresponding rotation angles. In this simulation the tool mass amounts to nearly 10% of the stage mass.

5 Control System Implementation

5.1 Hardware

The general structure of the control system shown in figure 4 is based on a dSPACE DSP system and no longer on a net of 7 transputers as reported in [4]. The greater number of analog outputs also provides for the individual control of all 8 coils of the drive. This enables a completely symmetric structure of the control software in particular for the generation of the torque that controls the drive's angular position. Also the comparatively weak power amplifiers based on power operational amplifiers were replaced by commercially available switched power amplifiers. They provide now $\pm 20A$ with a significantly reduced power loss.

5.2 Software

The control software used previously on the transputer hardware had to be rewritten completely. Fortunately, the generation of control software for the dSPACE system is quite comfortably supported by the MATLAB/SIMULINK extension REAL-TIME-WORKSHOP and the dSPACE REAL-TIME-INTERFACE on top of it which enable off-line simulation and testing as well as automatic code generation for the on-line program. The only alternative would have been a time consuming and error prone manual port of the OCCAM program to the programming language C. Therefore, it was attempted to implement all functions of the control system via the graphical interface. Figure 5 shows the hierarchical structure of the software.

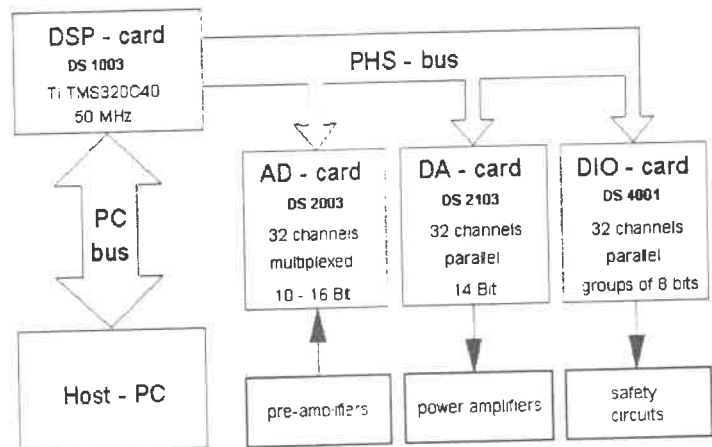


Figure 4. Hardware structure

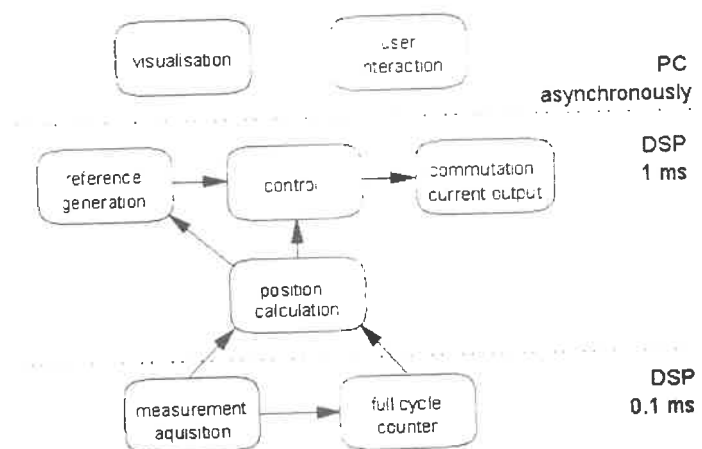


Figure 5. Software structure

The control system is implemented via the graphical interface. Figure 5 shows the hierarchical structure of the software.

6 Conclusion

Mechatronic analysis and design of high-precision drives can be done by integrating extended multi-body simulation with advanced symbolic control systems design implementing it on a rapid control systems prototyping platform. This approach results in short development cycles and paves the way towards an integrated design environment for mechatronic systems.

References

- [1] E. Kallenbach, K. König, E. Saffert, C. Schäffel, M. Eccarius. Integrated Design of Shape and Function in Mechatronic Systems. Proc. 3rd Conf. Mechatronics and Robotics, Oct. 1995 Paderborn, pp. 44-51.
- [2] O. Enge, G. Kielau, P. Maißer. Modelling and Simulation of Discrete Electromechanical Systems. Proc. 3rd Conf. Mechatronics and Robotics, Oct. 1995 Paderborn, Germany, pp. 302-318.
- [3] C. Schäffel, E. Saffert, E. Kallenbach. Integrated Electrodynamical Multicoordinate Drives - Modern Components for Intelligent Motions. Proc. ASPE 1996 Annual Meeting, pp. 456-461.
- [4] E. Saffert, C. Schäffel, E. Kallenbach. Control of an Integrated Multi-coordinate Drive. Proc. Mechatronics '96, Sept. 1996, Guimaraes, Portugal, vol.1 pp.151-156.
- [5] R. Rutz, J. Richert. CAMEL - An Open CACSD Environment, IEEE Control Systems Magazine, April 1995.
- [6] R. Kasper, J. Lückel, K.P. Jäker, J. Schröer. CACE Tool for Multi-Input, Multi-Output Systems using a New Vector Optimization Method. Int. Journal of Control 51 (1990), 5, pp. 963-993.

CURRENT AND FUTURE POSSIBILITIES OF NON-CONTACT NANO MACHINING BY COMBINATION OF LASER WITH SCANNING PROBE TECHNOLOGY

K. Dickmann, J. Jersch

University of Applied Sciences FH Muenster, FB Physikalische Technik, Laser Laboratory,
D – 48565 Steinfurt/GERMANY

Nano technology is of increasing interest for future oriented applications and bio technology. For the industrial production of nano components new nano handling systems as well as appropriate nano machining processes will be required. Various nano machining processes based on near field technology – known from Scanning Probe Microscopy (SPM) – have been developed in the past and will be introduced briefly in the following. However, this paper will mainly concentrate on a new laser machining process. FOLANT-process (Focusing of Laserradiation in the Nearfield of a Tip) is characterized by a couple of unique machining features and enables machining down to a few nanometers.

1. Introduction and current nano machining tools

Nowadays the potential of ion-beams, X-rays and electron beams for nano machining is well known. Another promising branch is based on SPM-technology. For example a simple known process utilizes the probe tip of a Scanning Tunneling Microscope (STM) for machining of nano structures by scratching into the material surface. Another contact nano machining process called "thermo mechanical writing" uses a laser heated tip from a Scanning Force Microscope (SFM). Nano structures for application in data storage are produced in polycarbonate by pressing the tip into the material. For industrial applications there is an increasing interest for non-contact nano machining processes. Thus, for example it is possible to create nano structures on metals when a short voltage peak is superimposed onto the tunneling voltage of an STM. It is also possible to modify magnetic surfaces with a resolution of some nanometers using magnetized probe tips. One of the optical nano machining methods is based on Scanning Nearfield Optical Microscopy (SNOM). Here laser radiation is guided through an optical fiber with an output aperture of some 10 nanometers onto the work piece. Nanostructures down to 50 nm can be obtained, when the fibre aperture is in the nearfield of the surface. Unfortunately, due to a reduced propagation of light, intensities are rather low and thus only machining of soft materials is possible. The listed disadvantages can be overcome by FOLANT-nano machining with considerable increased intensity and resolution¹⁻⁴.

2. Focusing of laser radiation in the nearfield of a SPM tip (FOLANT)

Based on various electrostatic ("lightning rod") and electrodynamic ("plasmon excitation") effects a considerable intensity enhancement of laser radiation occurs in the nearfield of a probe tip, when the tip is lateral illuminated by a laser. This effect is already well known from Surface Enhancement Raman Scattering (SERS) and enables enhancement factors up to 10^6 mainly depending on the tip material and -geometry. Large enhancement is obtained from STM metal tips with a good conductivity (e.g. silver, gold) but still is sufficient for SFM silicon tips. The intensity enhancement also increases with the sharpness of the tip. Thus a curvature radius of < 100 nm is necessary for applications in nano machining of metals.

The intensity enhancement only occurs in a narrow zone of some nanometers (horizontal and vertical) around the probe tip. Thus, for nano machining it is necessary to approach the material surface in to the nearfield of the tip. However, this can be easily obtained by state of the art SPM-technology. Computer controlled SPMs with user friendly scanning software also simplify the machining of defined geometry.

3. Setup for nano machining

The entire experimental setup for nano machining is shown in Fig. 1. A commercial scanning probe device (STM, SFM) combined with a frequency-doubled Nd:YAG-laser ($\lambda = 532$ nm) was used. The pulse energy was < 0.1 mJ by a pulse duration of 5 ns. The optics were used for matching the polarization to the longitudinal tip axis and for prefocusing the radiation into the tip/sample-gap.

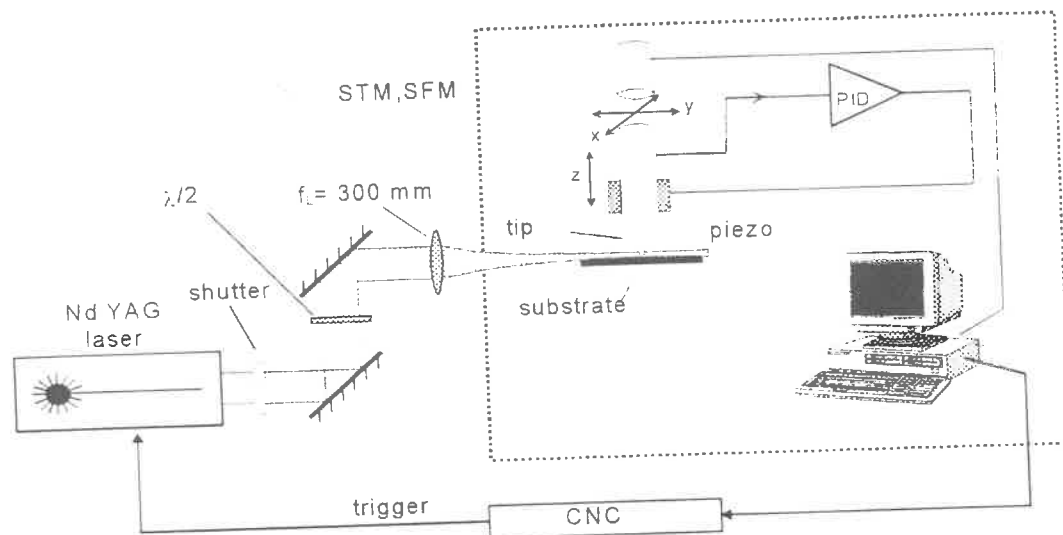


Fig. 1: Entire setup for nano machining in FOLANT configuration

4. Interaction processes and machining results

Various interaction processes of enhanced laser radiation with the substrate are possible, mainly depending on laser parameters (wavelength, pulse duration, pulse energy), substrate material and SPM-parameters (tip radius and -material, gap distance). We have determined thermal processes as well known from conventional laser material processing. We also believe in various non-thermal effects (e.g. "field evaporation").

The variety of process parameters optionally enable material removal from the substrate or deposition from the tip onto the substrate. The substrate variety ranges from metal to plastic. Thus nano-cutting, -drilling and -welding is possible. An overview of machining results is given in Fig. 2.

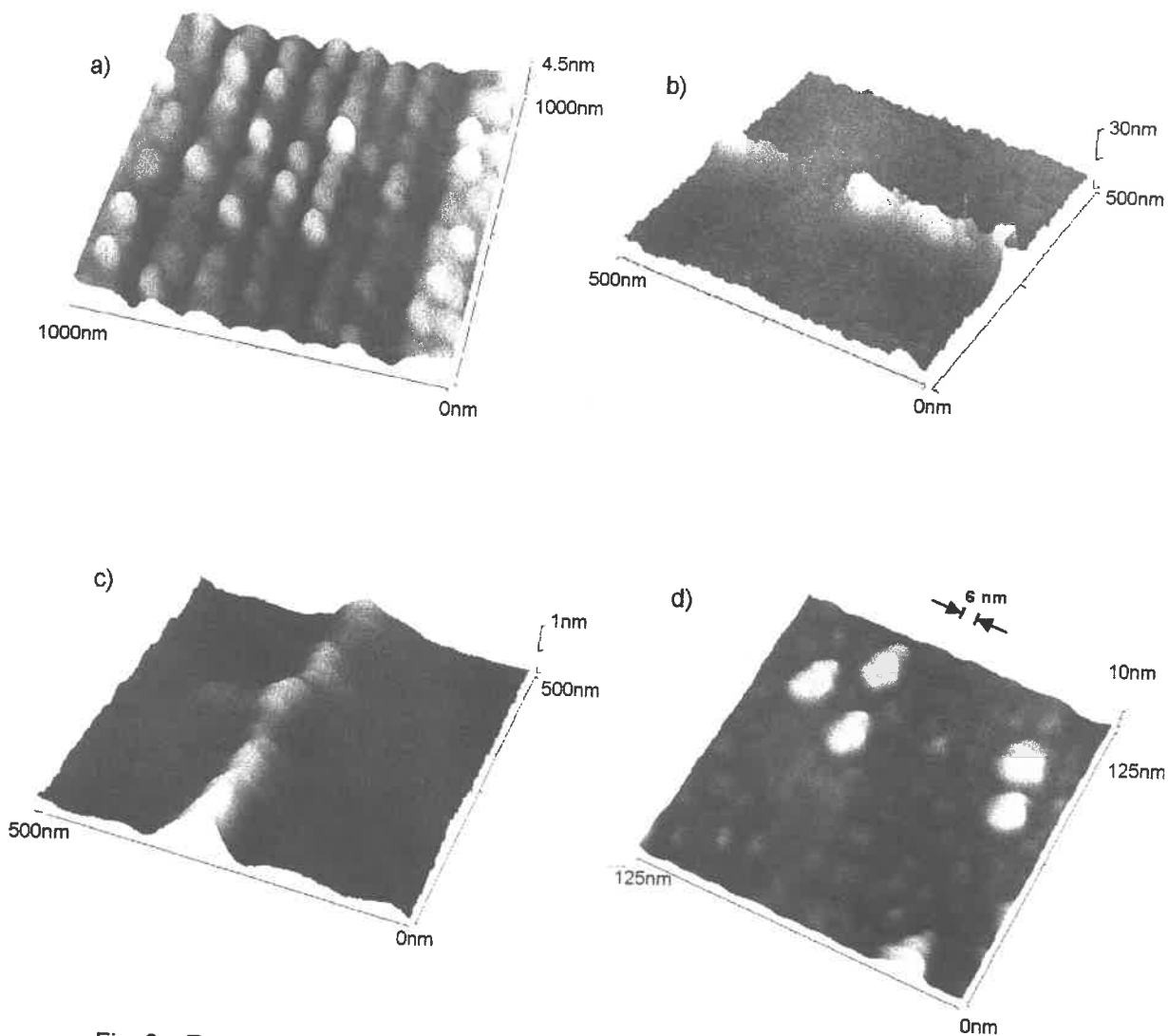


Fig. 2: Examples of nano machining

a) Parallel line cuts on a gold substrate

- b) Line structures on a gold substrate by material deposition for demonstration the potential in nano joining
- c) Nano structure on polycarbonate due to heating with a permanent surface modification
- d) Nano holes on a gold substrate (so far as known, these are the smallest structures ever achieved by la laser)

5. Outlook

The described method is a promising alternative to other known nano machining processes. The original purpose of the development was the application in data storage technique. With FOLANT technique already today we are able to increase the data density by 10^4 . Further increase up to 10^6 should be possible. Other applications can be found in mask repair of reticle for semiconductor industry, production of single electron transistors, production of X-ray fresnel optics and nano lithography. Since the introduced machining method is a basic technology, further applications will be found.

6. References

1. K. Dickmann, J. Jersch, Nanostrukturierung mit Laserstrahlung durch Feldverstärkung im Nahfeld einer RTM-Sondenspitze Laser und Optoelektronik 27 (3) 1995, p.76-83J.
2. J. Jersch, K. Dickmann, Nanostructure fabrication using laser field enhancement in the near field of a scanning tunneling microscope tip Appl. Phys. Lett. 68(6), 5 February 1996, p.868-870
3. K. Dickmann, J. Jersch, Nanostrukturierung mit Laserstrahlung Phys. Bl. 52 (1996) Nr.4. p. 363-365
4. J. Jersch, F. Demming, K. Dickmann, Nanostructuring with laser radiation in the nearfield of a tip from a scanning probe microscope Applied Physics A 64 (1), 29-32 (1997)

THERMAL EFFECTS OF BROAD BEAM NEUTRAL ION MACHINING

Helen E. Fawcett¹, Thomas G. Bifano², Donald Wroblewski³

Abstract

Broad beam neutral ion machining has been developed for a new manufacturing process for compact disc stampers: Precision Stamper Manufacturing (PSM). In this process, thermally induced degradation of the photoresist mask proved detrimental to the stamper quality, as can be seen from the following SEM micrographs.

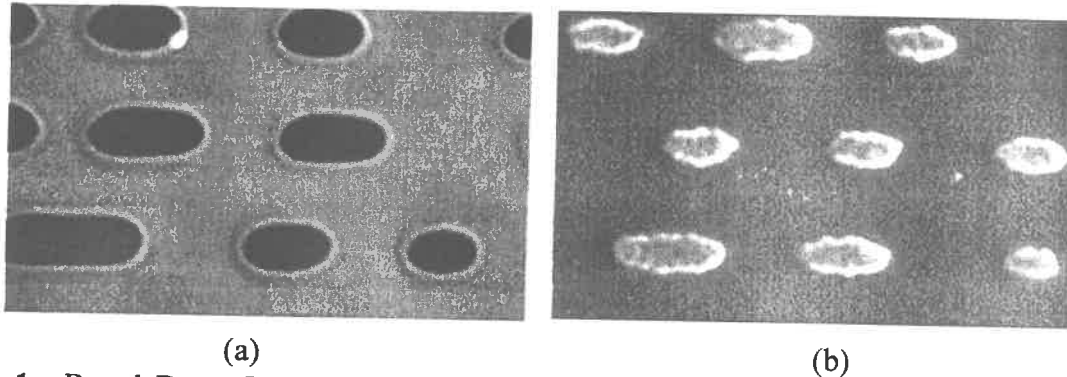


Figure 1. Broad Beam Ion Machining (680 nm thick photoresist): These SEM micrographs show the relationship between: (a), photoresist mask material on a CVD SiC substrate after LBR exposure and development, but before ion machining; and (b), the CVD SiC substrate surface after broad beam ion machining. (Magnification: 10,000 X)

Vacuum compatible thermocouples were placed in the vacuum chamber at various locations to record and analyze the thermal effects of the ion machining process. Results indicated that the ion beam (plasma) temperature exceeded 300°C in 30 minutes of machining. A silicon wafer was used as the substrate being etched during this test, and the temperature on the backside of the wafer exceeded the flow temperature of the negative tone photoresist (170°C) within the first few minutes of machining.

Thermocouples were imbedded into a copper cooling fixture to allow the measurement of various temperatures including the top and bottom plates and the substrate for determining heat transfer coefficients. From these experiments, a model of substrate temperature in the vacuum chamber during ion machining was developed. The model can easily be modified for different ion sources and machining conditions and should prove to be of general utility for thermally sensitive ion machining processes. This paper will present the results of ion machining compact disc features (0.6 μm wide) over a 138 mm diameter stamper without any form of cooling and corresponding results when cooling is added to the system. The model developed for this cooling fixture will also provide an opportunity to “predict” temperatures that a substrate will encounter during neutral ion machining.

¹ MicroE, 8 Erie Drive, Natick, MA 01760-1313 Tel: (508) 903-5023 Fax: (508) 903-5005 email: fawcett_helen@micro-e.com.

² Precision Engineering Research Lab, Boston University, 110 Cummington St., Boston, MA 02215 Tel: (617) 353-5619 Fax: (617) 353-5866 email: tgb@enga.bu.edu.

³ Aerospace and Mechanical Engineering Dept., Boston University, 110 Cummington St., Boston, MA 02215 Tel: (617) 353- Fax: (617) 353-5866 email: dew11@enga.bu.edu.

Introduction

The SEM micrographs shown in Figure 1 indicated that the photoresist had experienced thermal degradation during the ion machining process, therefore an initial experiment was conducted to quantify the thermal environment during machining. A shielded thermocouple was placed 1 inch above a silicon wafer that measured the plasma temperature (shown in Figure 2 as the thin line) and an exposed junction thermocouple was mounted under the wafer to measure substrate temperature (shown in Figure 2 as the thick line). It can be seen that without any cooling, the plasma and substrate temperatures during ion machining rapidly exceed the melting temperature of negative-tone photoresist (170°C).

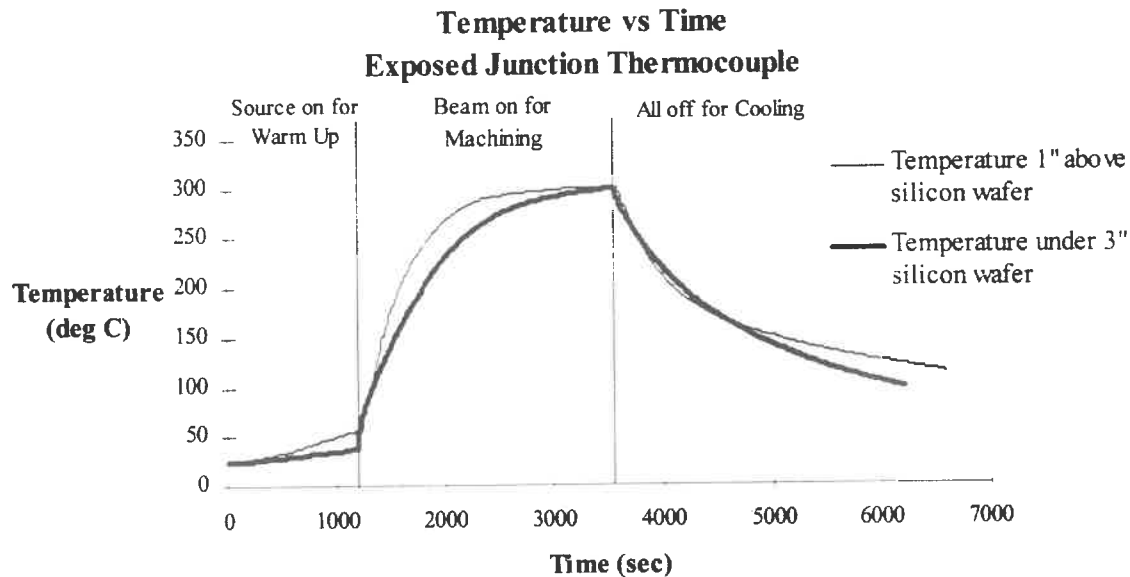


Figure 2. Thermal measurements 1 inch above and underneath a 3 inch silicon wafer.

Based upon the information shown in Figure 2, the need to characterize the substrate temperature before subjecting thermally sensitive material to the ion machining process became apparent. At the same time, the necessity of employing active cooling was evident in order to maintain the substrate temperature below 100°C ensuring that the photoresist mask would not degrade.

Heat Transfer Experiment

A lumped model describing the thermal energy flux during ion machining was developed. In concert with this modeling, a new copper chuck was designed and fabricated incorporating convective water cooling and allowing various temperatures to be measured for correlation with the thermal model. The chuck consisted of two parts: a copper top plate with integral flow channels to allow auxiliary convective cooling of the substrate, and a bottom plate of solid copper. A schematic of heat fluxes and stored energy in the substrate and the chuck is illustrated in Figure 3.

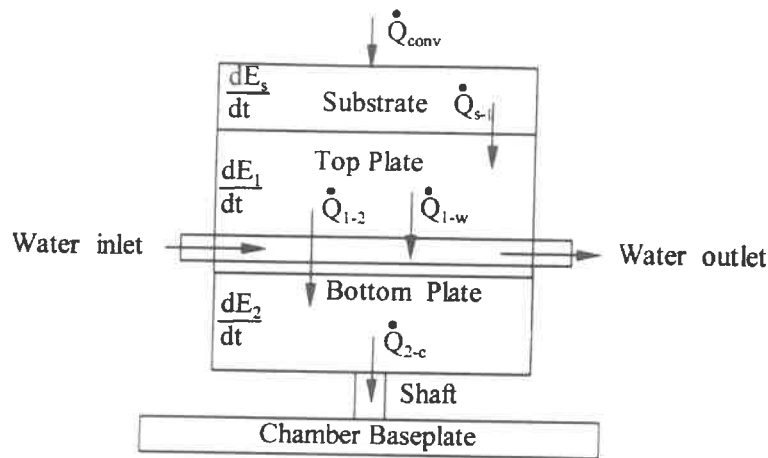


Figure 3. Schematic of thermal system for modeling analysis.

where $\dot{Q}_{conv}$ = convective heat flux from the ion plasma to the substrate

$\dot{Q}_{s-1}$ = heat flux from the substrate to the top plate (1)

$\dot{Q}_{1-w}$ = heat flux from the top plate (1) to the water

$\dot{Q}_{1-2}$ = heat flux from the top plate (1) to the bottom plate (2)

$\dot{Q}_{2-c}$ = heat flux from the bottom plate (2) to the chamber

$\frac{dE_s}{dt}$ = energy stored in substrate

$\frac{dE_1}{dt}$ = energy stored in top plate (1)

$\frac{dE_2}{dt}$ = energy stored in bottom plate (2)

The principal input of thermal energy for this system is convective heat flux from the Argon ion plasma to the substrate surface. A portion of this energy is stored in the substrate in the form of increased temperature. Heat flux out of the substrate is principally by conduction (through the substrate/chuck interface). The heat flux into the upper half of the chuck can be stored, convected to the water flowing through it, or conducted to the lower half of the chuck. Radiation from the substrate and chuck is assumed to be negligible based on a simple estimate of radiative heat flux using known bounds on substrate and chamber wall temperatures. A thermal energy balance on the substrate yields:

$$\dot{Q}_{conv} = \frac{dE_s}{dt} + \dot{Q}_{s-1} \quad (1)$$

The small thermal mass of substrate and its close coupling to the copper chuck allows stored heat in the substrate to be neglected.

Thus Equation 1 can be rewritten as:

$$h_i A_s (T_\infty - T_s) = \frac{(T_s - T_1)}{R_c} \quad (2)$$

where R_c = Thermal contact resistance at the substrate/chuck interface (K/W)

T_∞ = Plasma temperature (°C)

T_s = Substrate temperature (°C)

T_1 = Top plate temperature (°C)

A_s = Surface area of the substrate/chuck interface (m²)

h_i = coefficient of the plasma's convective heat transfer (W/m² - K)

While time dependent quantities, T_s , T_1 , and T_∞ , can be measured, heat transfer coefficients, h_i , and R_c , are unknown. To estimate these heat transfer coefficients, ion machining experiments were conducted without a substrate. The energy equation for the chuck alone during ion machining is:

$$\dot{Q}_{conv} = \frac{dE_1}{dt} + \frac{dE_2}{dt} + \dot{Q}_{2-c} + \dot{Q}_{l-w} \quad (3)$$

Once the heat transfer coefficients were evaluated, further simplifications to Equation 3 allowed an expression for the substrate temperature (T_s) to be obtained.

$$T_s \approx \frac{R_c h_i A_s T_\infty + T_{in}}{1 + R_c h_i A_s} \quad (4)$$

Equation 4 predicts T_s for steady state and is not a prediction for the time dependent temperature. It can be seen that the two critical parameters defining substrate temperature are interface contact resistance and the convective heat transfer coefficient. Only R_c can be modified substantially in the present system. Factors that decrease R_c and therefore decrease T_s include surface finish, contact pressure, and surface conditions. Also, interfacial vacuum compatible greases exist that increase thermal flux. Equation 4 required only three temperatures to be measured during an ion machining process, one of which is on the atmospheric side. The model can be used as an estimate of the temperature that an actual sample would encounter during ion machining.

Experiments were then conducted with a substrate on top of the cooling chuck. Figure 4 shows measured values for the substrate temperature (T_s), and the inlet water temperature (T_{in}): the plasma temperature (T_∞) was similar to that shown in Figure 2.

Temperature vs Time With Substrate, With Cooling

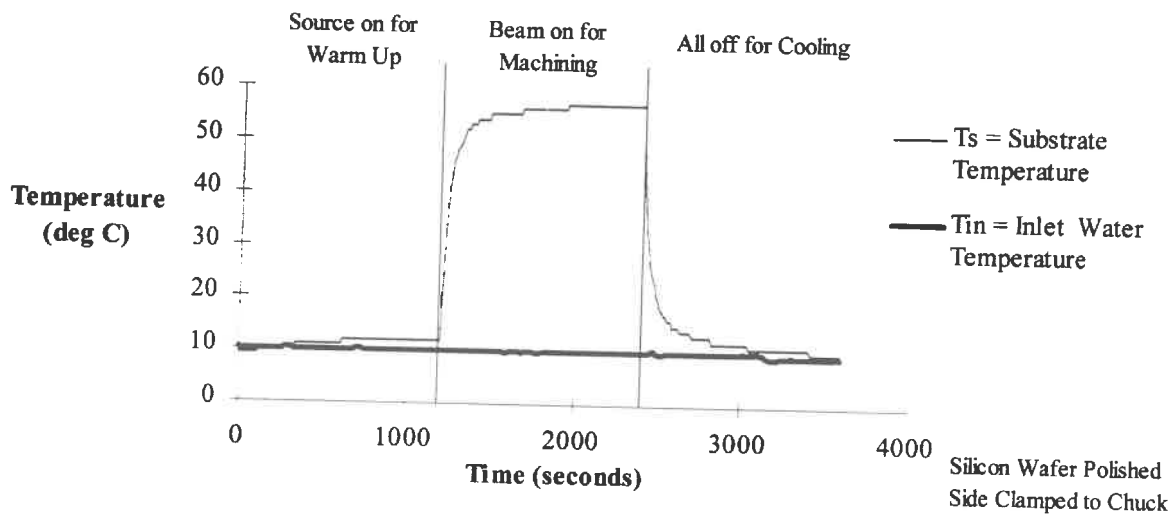


Figure 4. Temperature versus time for the plate with cooling and with a substrate on top. Shown in this figure are the temperatures of the inlet water temperature (T_{in}) and the substrate (T_s).

Machining Experiments

An electroless nickel stamper was used as the first to be machined with active cooling. SEM micrographs of the photoresist before and after the ion machining process with active cooling are shown in Figure 5.

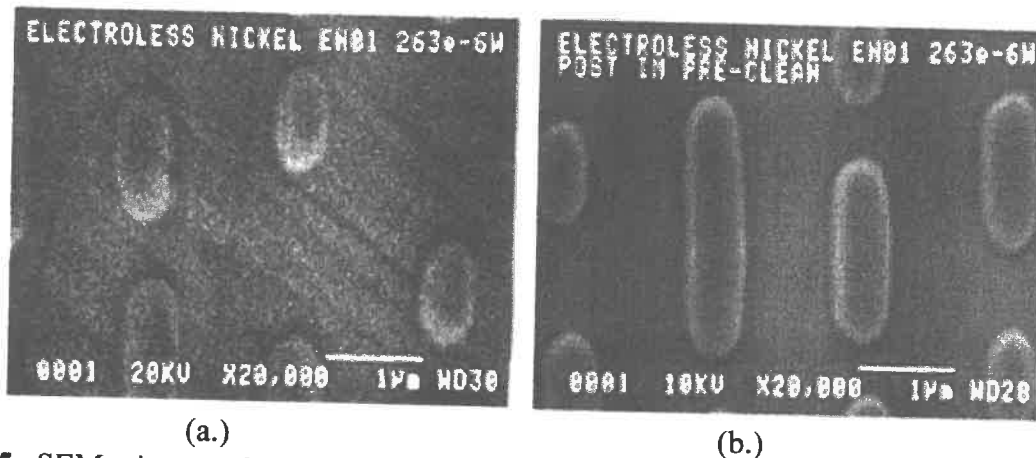


Figure 5. SEM micrographs of electroless nickel stamper. (a) photoresist exposed on electroless nickel stamper surface and (b) resulting feature on electroless nickel stamper after ion machining with active cooling. (Magnification: 20,000 X)

As can be seen from Figure 5, no distinguishable change in the CD feature integrity was observed. This stamper was then used to injection mold polycarbonate CD's from it in an attempt to characterize the quality of PSM compact discs when compared to those made through

attempt to characterize the quality of PSM compact discs when compared to those made through conventional processes. The following AFM results show comparable quality of feature transfer into the polycarbonate CD between PSM and conventional mastering processes.

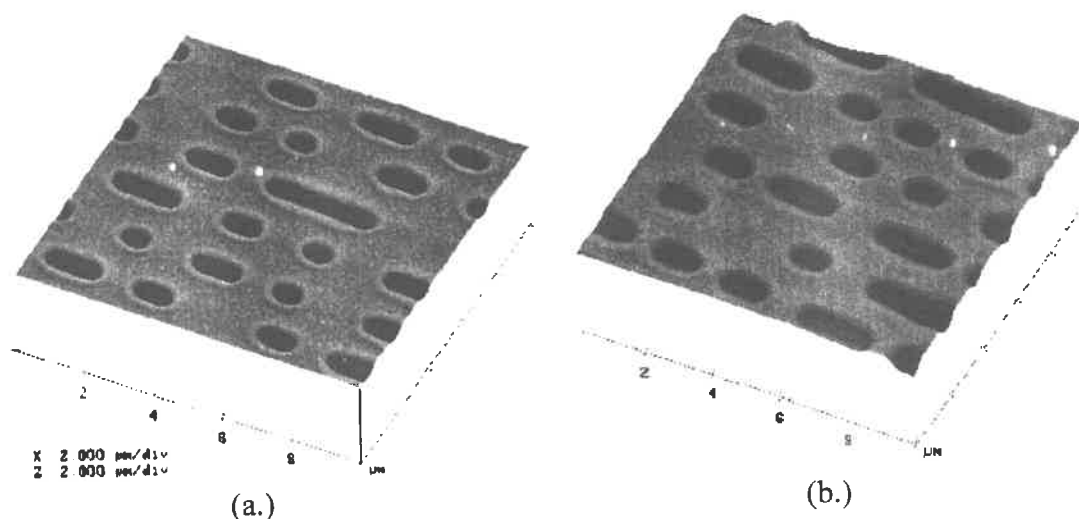


Figure 6. AFM Comparison of CD replicas in polycarbonate. (a.) shows a replica molded from a conventional stamper and (b.) shows a replica molded from an electroless nickel PSM stamper.

Conclusions

The model and cooling chuck designed in this research fulfilled the objective of maintaining the temperature of the substrate below 100°C. The information in Figure 4 also shows that throughout 15 minutes of ion machining, the temperature of the substrate did not exceed 60°C. Since the machining time of the PSM process was limited to approximately 5 minutes, the thermal conditions were considered to be adequate for the process. The ability to machine stampers using the PSM process without the photoresist mask melting was achieved and the results of the thermal model and corresponding machining of the stampers will be presented.

Acknowledgments

This research was funded in part through the National Science Foundation Grant DMI 9500162 and Prism Corporation. We also wanted to thank the following people for their contributions to this research. Anlee Krupp who completed all of the SEM work, Steven Fawcett who worked on the design of the cooling chuck, Margo Velonis and all members of BU-PERL who helped with the thermal measurements, and Ion Tech, Inc for all their technical support.

References

- [1] Fawcett, Helen E., May 1997, "Precision Manufacturing of Optical Storage Stampers," Ph. D. Dissertation.
- [2] Lienhard, John H., 1981, A Heat Transfer Textbook: Second Edition. Prentice Hall, Inc., New Jersey.
- [3] Melliar-Smith, C. M., "Ion etching for pattern delineation." *Journal of Vacuum Science and Technology*, Vol. 13, No. 5, p. 1008 - 1022.
- [4] Pohlmann, Ken C., 1992, Volume 5: The Computer Music and Digital Audio Series: The Compact Disc Handbook 2nd Edition. A-R Editions, Inc., Wisconsin.

Near-field Photolithography to Realize High Resolution Smaller Than Light Wavelength ~ Stamp Photolithography

Shuji Tanaka*, Masayuki Nakao and Yotaro Hatamura
Department of Engineering Synthesis, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113 Japan

1 Introduction

Lithographic techniques for micro device fabrication require high resolution, low cost, large throughput, high yield, accurate registration and so on.

A noncontact photolithographic technique has the advantages of large throughput, high yield and accurate registration, so that it is used mainly for the production of semiconductor integrated circuits. This technique, however, has the limit in resolution determined by the wavelength of exposure light because of diffraction limit. In order to realize higher resolution, shorter wavelength of exposure light is required. Patterning features with sizes less than $0.2\ \mu\text{m}$ is, however, difficult even if deep ultraviolet light with wavelength of about $200\ \text{nm}$ is used. Furthermore, a costly photolithographic device with a deep ultraviolet light source and an optical system for it results in large equipment investment and high production cost.

A conventional contact photolithographic technique is free from diffraction between a mask and a photoresist, and can realize high resolution smaller than the wavelength of exposure light. High yield is, however, difficult to realize without costly exchange or cleaning of a mask because of the damage to the mask caused by the contact with a photoresist. This technique is not also suited for registration, because the contact between a mask and a nonplanar surface is difficult.

An electron beam lithographic technique has the advantages of high resolution, high yield and accurate registration, while it also has the significant problems of throughput and cost.

As a photolithographic technique to complement the conventional ones, the authors propose a new near-field photolithographic technique to realize high resolution, large throughput and high yield at **low cost**. This technique, as Figure 1 shows, comprises the process of electron beam lithography to fabricate a master mold with a relief, replication of the master mold with a transparent replica material and near-field exposure with the replica mold, followed by etching or lift-off process.

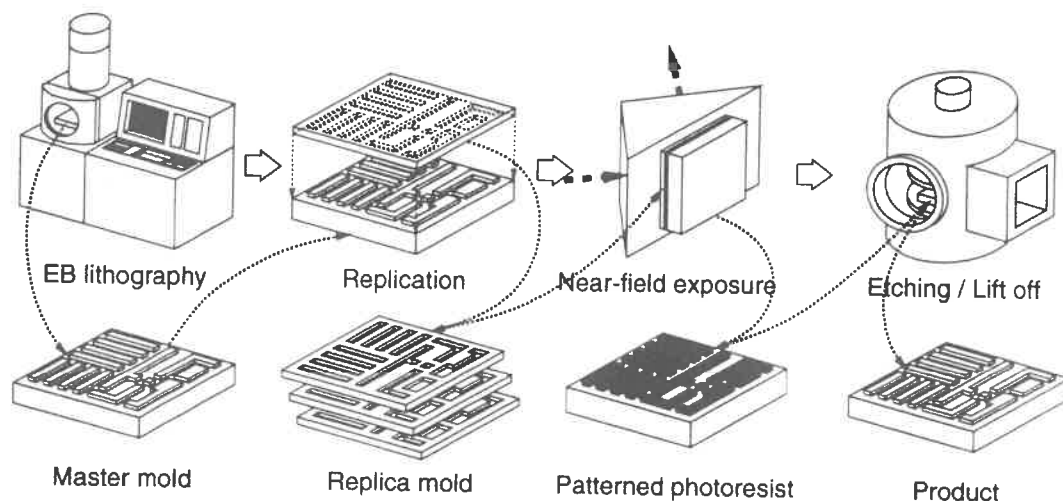


Figure 1 Process of the near-field photolithography

Figure 2 illustrates the method of the near-field exposure. A photoresist is attached or approximated to a relief on a transparent mold. The mold is illuminated through a prism in the condition of total internal reflection. The photoresist is exposed mainly to near-field light transmitted from the convex part of the

*Supported by JSPS Research Fellowships for Young Scientists

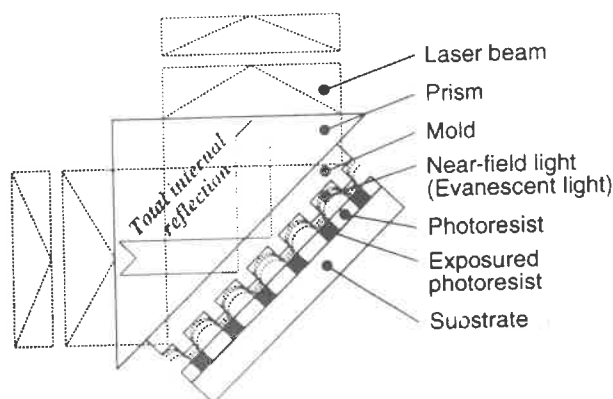


Figure 2 Method of the near-field exposure

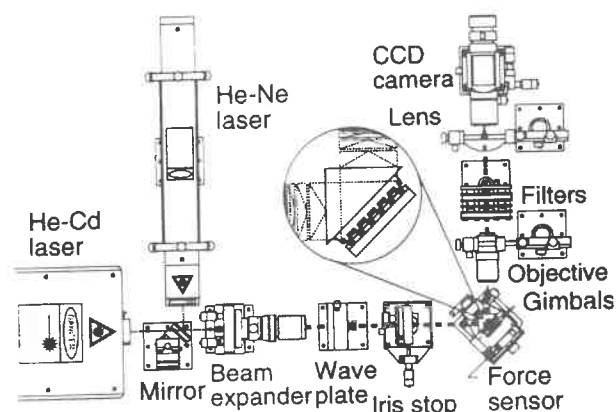


Figure 3 Schematic of the experimental device for the near-field exposure

relief. The authors call this technique "**Stamp Photolithography**", because the method of the exposure is like stamping.

The stamp photolithography can realize high resolution smaller than the wavelength of exposure light, because the exposure is performed in the near-field. A replica of a master mold fabricated at low cost serves as a disposable replica mold, which circumvents the reduction of yield caused by the damage to the mold at low cost. Furthermore, this technique does not require a light source with short wavelength or a high-performance optical system, so that the photolithographic device is inexpensive and a conventional photoresist can be used.

2 Experiment

An 80 μm thick replica film of acetyl cellulose is used for replication of a master mold with a relief fabricated by electron beam lithography. Figure 3 shows an experimental device for the near-field exposure. It is equipped with a He-Cd laser with wavelength of 441.6 nm (rated output: 10 mW, beam diameter: $\phi 0.9$ mm, TEM₀₀, linear polarization) to expose a G-line (wavelength: 446 nm) photoresist, and a He-Ne laser with wavelength of 632.8 nm to observe the position of a relief on the mold and the contact condition between a mold and a photoresist. The laser beam is expanded 10 times as large in diameter by a beam expander, polarized by a $1/2\lambda$ wave plate, narrowed down by an iris stop, and totally reflected in a prism with a replica mold on the slanting surface. A substrate with a spin-coated photoresist is fixed on a 0.2 mm thick gimbal plate, whose self-alignment mechanism allows the close contact between the mold and the photoresist. The contact force is measured by a force sensor (rated load: 50 N). The laser beam totally reflected in the prism is led to a CCD camera through an objective, filters and a lens. The CCD camera allows observing the position of the relief on the mold and the contact condition between the mold and the photoresist.

The authors conducted experiments along the process shown in Figure 1. First, the authors fabricated a master mold of silica glass with a relief of gratings with pitch of 100 nm $\sim$ 1 μm (width of 50 nm $\sim$ 500 nm) by electron beam lithography and fast atom beam etching. The depth of the relief is about 60 nm. Figure 4 shows an SEM view of the master mold.

Secondly, the authors replicated the master mold with a replica film of acetyl cellulose. The replica is easy to fabricate only by attaching the replica film dipped into methyl acetate to the master mold. Figure 5 shows an AFM view (a top view and a cross section) of the replica mold. The depth of the relief on the replica mold becomes shallower as the pitch of the grating decreases: the depth of the relief of the grating with pitch of 1 μm is about 50 nm, and that with pitch of 360 nm is about 15 nm.

Lastly, the authors attached the replica mold to the prism so that the grating is set parallel to the laser direction, put a Si substrate (about 5 mm square) with a spin-coated G-line positive photoresist (thickness: about 100 nm) against the replica mold by the contact force of 5 N observing the contact condition with the He-Ne laser and the CCD camera, and exposed the photoresist for 7 seconds. Figure 6 shows an optical microscope view of a pattern on the photoresist developed for 65 seconds at room temperature. Figure 7 shows an AFM view of the framed parts in Figure 6, and it corresponds with the AFM view of the replica mold in Figure 5. The grating pattern with pitch of 500 nm $\sim$ 1 μm (width of 250 nm $\sim$ 500 nm) was clearly printed, and that with pitch of 360 nm $\sim$ 400 nm (width of 180 nm $\sim$ 200 nm) was also printed, but

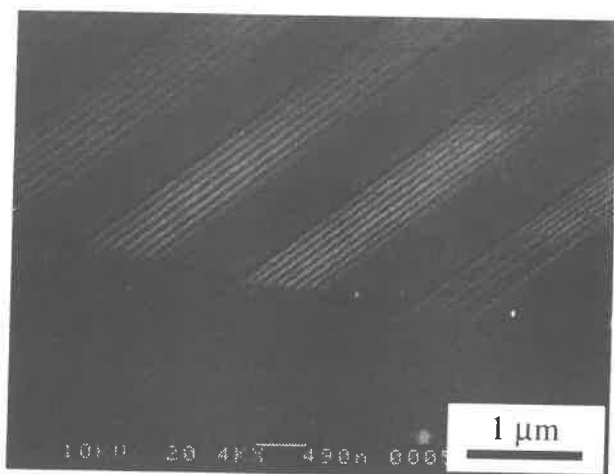


Figure 4 SEM view of the master mold

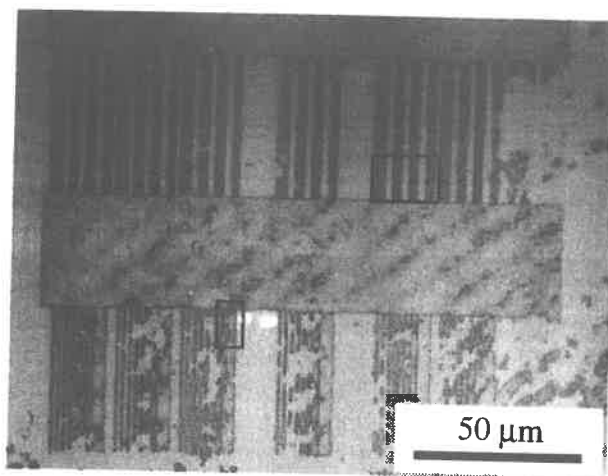


Figure 6 Optical microscope view of the pattern on the photoresist

not as clearly as in the first case. The grating pattern with pitch less than 360 nm was not printed in this experiment. The laser beam was p-polarized, because s-polarized laser beam extremely reduces the contrast of a printed pattern.

3 Discussion

The important conditions of the near-field exposure include the depth of a relief on a (replica) mold and the contact condition between a mold and a photoresist in addition to exposure time and polarization.

A deep relief on a mold enhances scattering light from the sides of the relief and lengthens the distance from the scattering light source to a photoresist, so that the scattering light interferes and many interference fringes are projected on to the photoresist. In another experiment with a mold of silica glass with a 700 nm deep relief, many interference fringes were observed on a photoresist [1].

A shallow relief on a mold, on the contrary, is elastically deformed and flattened by the contact with a photoresist. It is one of the probable causes that the grating pattern with pitch less than 360 nm was not printed in this experiment.

As for the contact condition between a mold and a photoresist, the close contact and the minimum deformation of a relief on a mold are necessary. It is thought that a replica material which allows a conformal contact with a photoresist without being pressurized to a photoresist can achieve a good contact between the mold and the photoresist.

Now the authors are fabricating new master molds and trying new replication methods to realize high resolution of 200 nm ~ 300 nm.

4 Conclusion

The authors proposed the stamp photolithography to realize high resolution, large throughput and high yield at low cost. The grating pattern with pitch of 360 nm (width of 180 nm) was printed using exposure light with wavelength of 441.6 nm, and the high resolution smaller than the wavelength of exposure light was achieved. Future work will be to improve the feature of a relief on a mold and a replica material to realize higher resolution.

The authors would like to thank Dr. Masahiro Hatakeyama and Mr. Yasushi Tohma at Ebara Research Co., Ltd. for their assistance in electron beam lithography and fast atom beam etching. The authors also would like to thank Mr. Hideki Miyazaki at Research Center for Advanced Science and Technology, The University of Tokyo for his helpful advice.

References

- [1] Shuji Tanaka, et al., Experiment of Simultaneous Patterning by Evanescent Light Lithography, Proceedings of The 74th JSME Spring Annual Meeting, 1997. Vol. IV, pp. 322-272

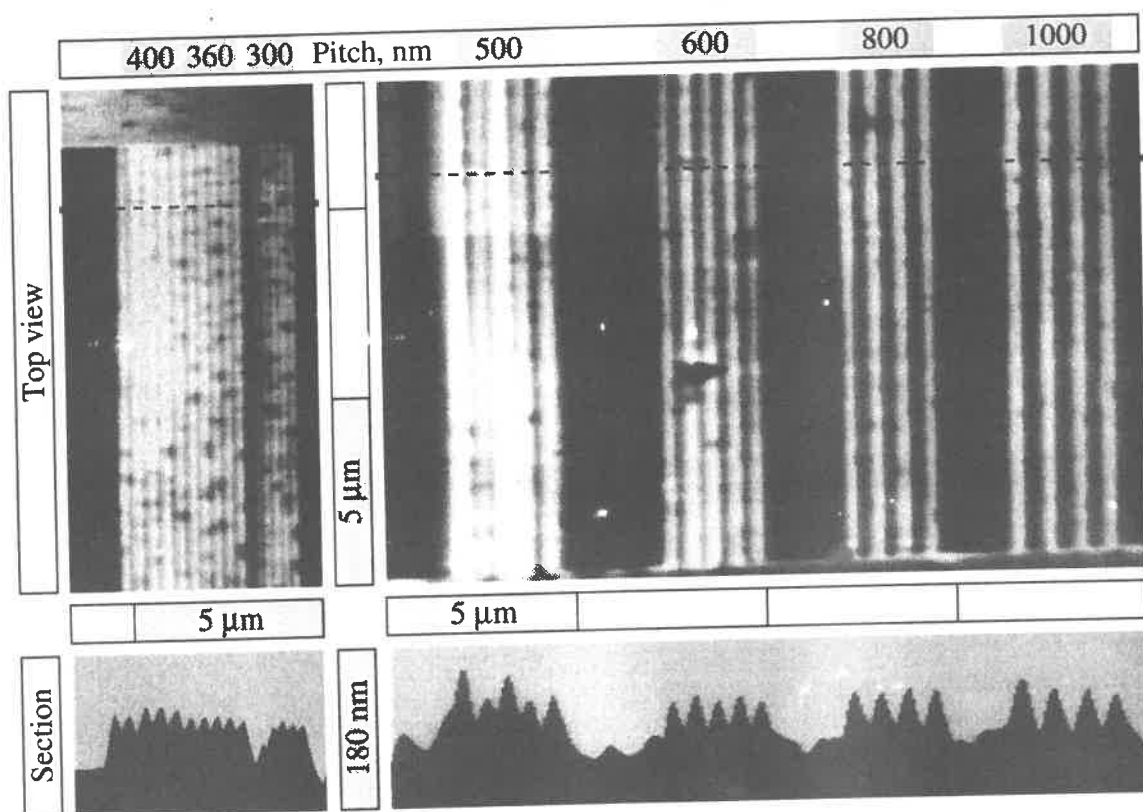


Figure 5 AFM view of the replica mold (a top view and a cross section)

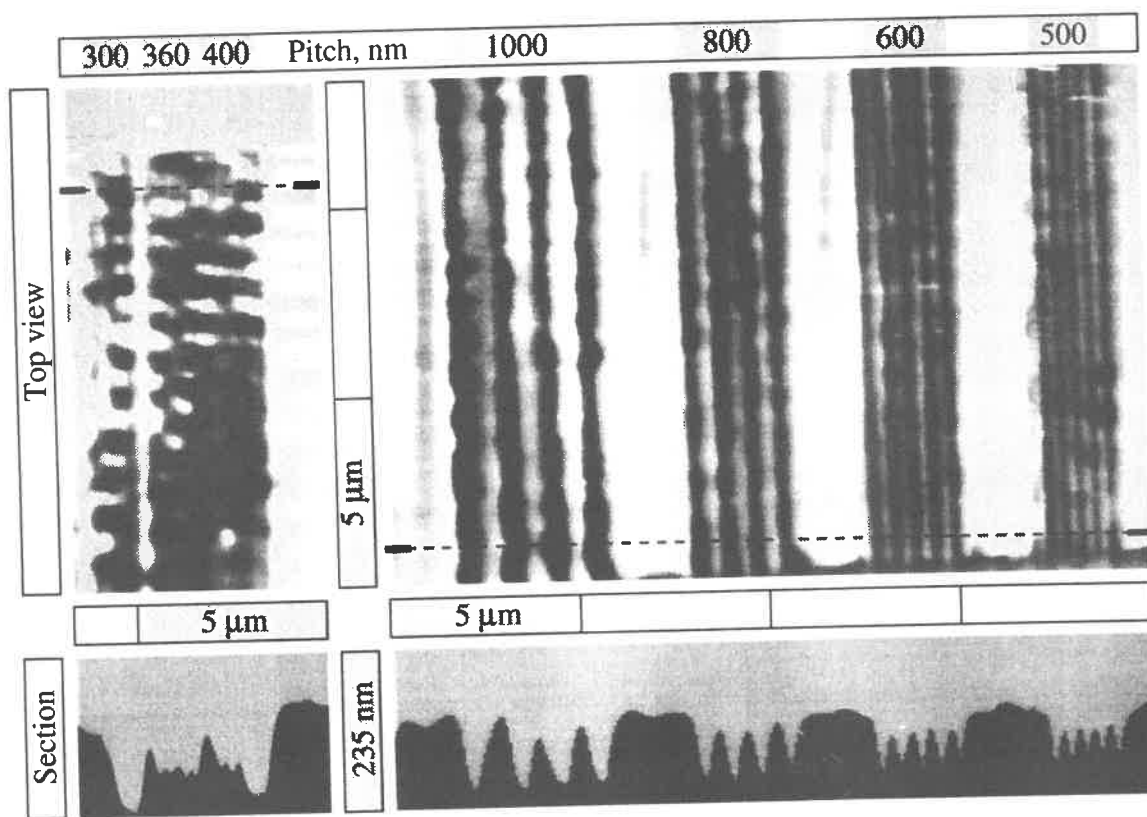


Figure 7 AFM view of the pattern on the photoresist (a top view and a cross section)

High Efficiency Fine Boring of Single Crystalline Silicon Ingot by Electrical Discharge Machining

Yoshiyuki UNO*, Akira OKADA*, Yasuhiro OKAMOTO* and Yoshiaki YAMADA**

* Okayama University, 3-1-1, Tsushimanaka, Okayama 700, JAPAN

** KASEN USA Corp., 9101 Southern Pines Blvd., Charlotte, NC 28273, USA

1. Introduction

Single crystalline silicon is one of the most important materials in the semiconductor industry because of its many excellent properties as semiconductor. In the manufacturing process of integrated circuits, plasma etching process is used for removing the oxidation film such as SiO_2 .¹⁾ In this process, graphite plate with many fine holes has been used as an electrode. Recently, it is examined to use single crystalline silicon as the electrode in order to minimize the contamination. Diamond grinding quill has generally been used for boring of silicon so far. However, it is difficult to machine accurately by the conventional methods, since the material removal is due to brittle fracture. In the diamond drilling method, the removal rate decreases and the loading of tool becomes severe with increasing of the machining depth. Furthermore, the chipping of workpiece occurs at the exit of hole. Therefore the usage of electrical discharge machining (EDM) is proposed for boring fine holes of single crystalline silicon. Then the possibility of high efficiency and high accuracy boring holes in silicon ingot by EDM is experimentally investigated.

2. Experimental Procedures

Fig.1 shows the schematic diagram of experimental apparatus. A copper pipe of 1 mm in diameter is used as an electrode, which rotates at 90 rpm. The unflammable type fluid is spouted out from the tip of the electrode. Single crystalline silicon ingot used as an electrode in plasma etching process has low resistivity in the order of $0.01 \Omega \cdot \text{cm}$, which makes it possible to machine silicon by EDM. Mild steel SS400 in JIS specification is also employed as a workpiece for comparison. Laser displacement sensor is set for in-process measurement of the electrode movement. The machining conditions are shown in Table 1. The polarity of electrode is kept plus through the experiment because the removal rate is higher than the reverse case.

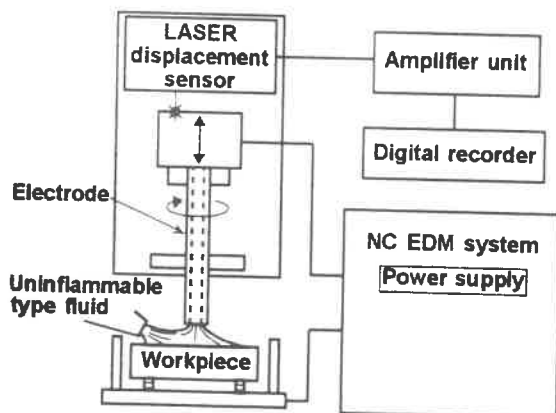


Fig.1 Schematic diagram of experimental apparatus

Table 1 Machining conditions

Polarity	Electrode(+)
Discharge current	$I_p=2,6,15\text{A}$
Pulse duration	$\tau_p=4,8,12,20.28 \mu\text{s}$
Duty factor	D.F.=50%
Capacity	$C=380\text{pF}$
Machining fluid	Unflammable type

3. Results and Discussions

Fig.2 shows the relationships between the removal rate and the pulse duration in EDMing of silicon and mild steel. Here, the removal rate is expressed as the ratio of workpiece thickness to total machining time. As shown in the figure, the removal rate of silicon is about 4 to 8 times larger than that of steel when compared under the same machining conditions. In EDM, it is generally proved that the removal rate is controlled by the product of melting point θ_m and heat conductivity λ , and low value of $\theta_m \cdot \lambda$ leads to high removal rate²⁾. The value of $\theta_m \cdot \lambda$ for silicon is $14.1 \times 10^5 \text{ W/m}$, while that for steel is $1.04 \times 10^5 \text{ W/m}$. From this point of view, the removal rate of steel is expected much larger than that of silicon. However, the experimental result shows opposite tendency. It is recently reported by Saeki et al. that the Joule's heat generation played an important role in material removal in EDM of high resistance material³⁾. Also in this case, it is considered that the Joule's heat generation in addition to heat conduction from arc column makes a great contribute to increasing the removal rate of silicon. On the other hand, the electrode wear of silicon is 1/15 to 1/7 compared with that of steel under the same conditions. Therefore, EDM is a very efficient method to machine single crystalline silicon.

Fig.3 shows an example of the displacement of electrode during machining time in EDMing of silicon. The machining process is divided into three distinct regions with the machining time: normal region, stagnation region and quick feed region. In normal region, the machining is stable and the machined depth increases linearly with the machining time. When the electrode advances around the exit of hole, the stagnation region appears. In this region, the electrode retracts suddenly and the machining becomes instable. After the stagnation region, the boring finished and the electrode advances at the rapid feed speed. **Fig.4** shows the SEM micrographs of the cross sections of machined hole. Good machined surface is generated around the entrance of the hole, while sticking of the insulator is observed on the wall of hole around the exit. The EPMA analysis pointed it out that the insulator contained silicon oxide and copper oxide. **Fig.5** is the explanation diagrams of these phenomena. As shown in (a), the exclusion of debris is smoothly performed through the gap between the workpiece and the electrode until the hole is penetrated. However, once the hole is penetrated as shown in (b), the exclusion of debris becomes difficult and they stagnate around the exit because the machining fluid leaks through the hole. The stagnation of debris causes the secondary discharge between the debris and the workpiece, and it makes the insulator stick on the wall of hole.

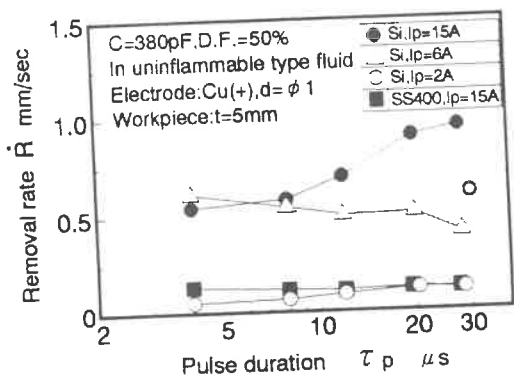


Fig.2 Relationship between removal rate and pulse duration

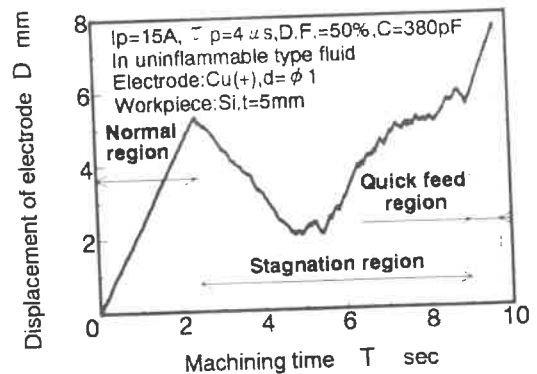


Fig.3 Displacement of electrode with machining time

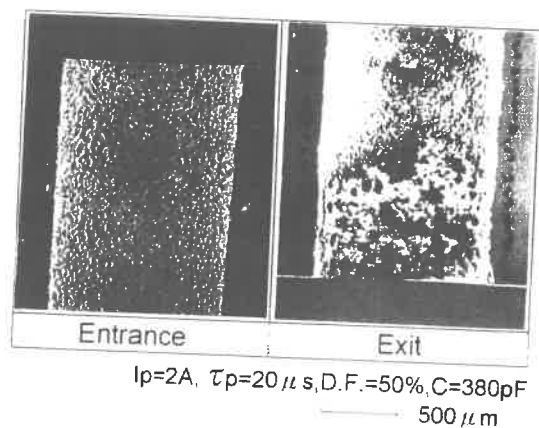


Fig.4 Cross sections of machined hole

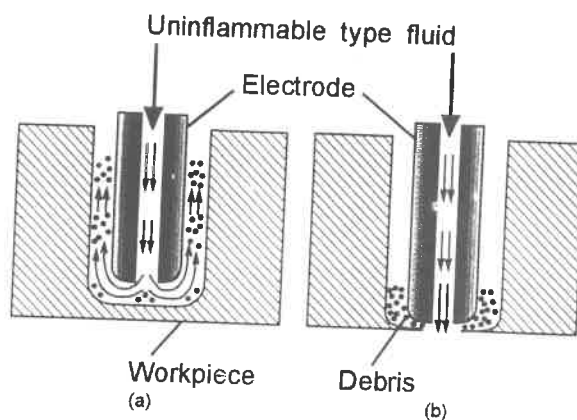


Fig.5 Exclusion of debris

4. Improvement Method of Machining Around Exit

In order to maintain the flow of the machining fluid well and make the exclusion of debris smooth even around the exit of hole, the improvement method is introduced as shown in Fig.6. In this method, the copper plate is attached under the silicon firmly. Fig.7 shows the micrographs of machined hole generated with this method. Better hole can be obtained without sticking of the insulator and chipping around the exit. Fig.8 shows the relationship between the displacement of electrode and the machining time. There is no stagnation region, which leads to better machined hole.

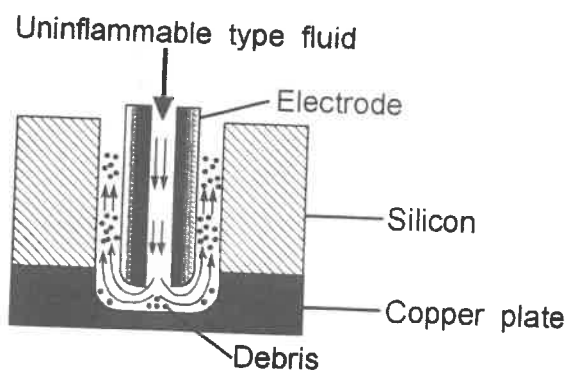


Fig.6 Improvement method of machining around the exit of hole

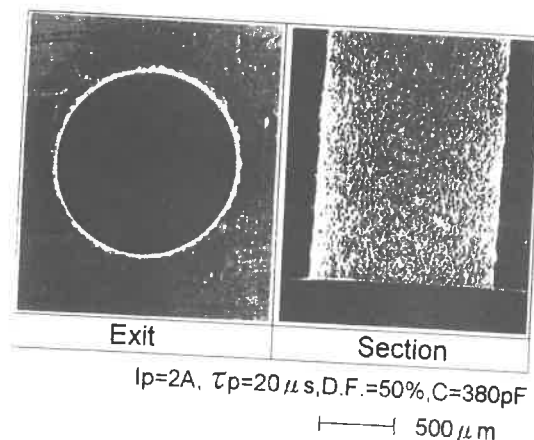


Fig.7 SEM micrographs of machined hole by improvement method

5. High Aspect Ratio Boring with EDM

Thicker electrode of silicon for plasma etching is required in order to reduce the arrangement time. We tried the deep boring of 200mm silicon which is the diameter length of 8" ingot. Fig.9 shows the relationship between the displacement of electrode and the machining time.

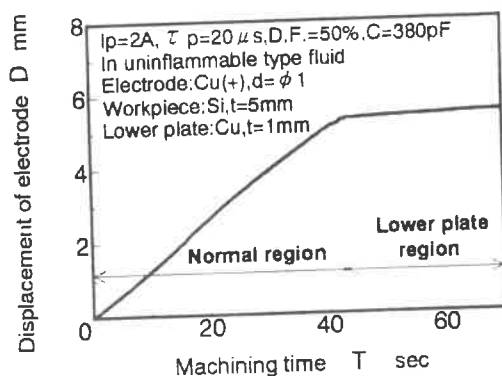


Fig.8 Relationship between displacement of electrode and machining time by improvement method

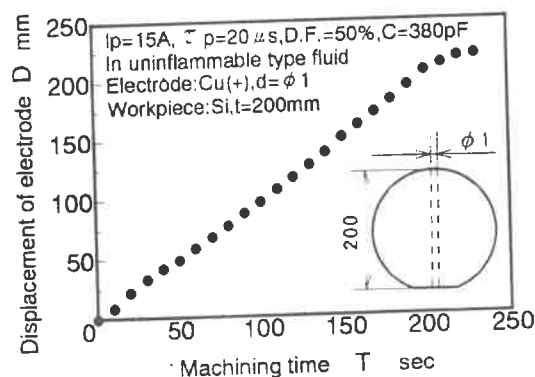


Fig.9 Relationship between displacement of electrode and machining time in high aspect ratio boring

Under this condition, stable machining is performed at the rate of 1mm per second and the straightness is maintained well. Judging from the experimental result, it seems that deeper hole than 200mm is possible to machine with this method. Generally speaking, it becomes difficult to machine in boring of deep hole because the loading of drill or the obstruction of debris with the machine in boring of deep hole. However, it progresses at the constant speed and it is possible to bore the deep hole whose aspect ratio is about 200 in this method. From the above results, it proved that EDM is the suitable method for boring of single crystalline silicon.

6. Conclusions

Main conclusions obtained in this study are as follows:

- (1) The removal rate of single crystalline silicon is much higher than that of steel, while the electrode wear is smaller.
- (2) The exclusion of debris is difficult around the exit of hole, which leads to the inefficient stagnation region and the sticking of insulator made from silicon oxide and carbon oxide onto the wall of hole.
- (3) The improvement method in which the copper plate is attached under the silicon firmly leads to better results.
- (4) High aspect ratio boring of single crystalline silicon is efficiently attained by EDM.

Acknowledgment

The authors would like to express their thanks to Sin-Etsu Handotai Co. for supplying the precious workpieces of single crystalline silicon.

References

- 1) T.SUGANO: Plasma Process Technology for Semiconductor, SANGYO-TOSHO(1993)228.
- 2) N.SAITO: Principle of EDM and its Application, GIJUTSU-HYORONSHA(1994)38.
- 3) T.SAEKI et al.: Transient Workpiece Temperature Analysis in the EDM Processes of High Electric Resistance Materials Considering Joule's Heating, J. of JSPE, 62,3(1996)443.

Micro Devices Produced by Miniature Robots

*Hisayuki Aoyama, **Ken'ichiro Mizutani, **Tadahiro Miura,
#Kenji Murakami and **Futoshi Iwata

**Applied Micro Systems of Robotics Engineering
Dept. of Mechanical & Control Engineering
University of Electro-Communications
1-5-1, Chofu, Tokyo 182, Japan*

***Dept. of Mechanical Engineering,
#Research Institute of Electronics
Shizuoka University
3-5-1, Hamamatsu, Shizuoka 432, Japan.*

1. Introduction

Recently many new approaches have been developed to make the micro parts with the help of MEMS technology. One of them is an ultra precise machining techniques based on the mechanical tooling and another is on the electro-chemical technique which is widely used in LSI process. From the view point of micro-material processing, the former can be represented as a "top-down" approach involving material removal and the later is as a "bottom-up" approach with material deposition and etching. Such an ultra precise and a micro fabrication in micro realm need the multi-disciplinary experiences of several novel processes that can create complex micro structures.

In this report, many miniature robots of the beetle size with micro tools and workable in a vacuum chamber for the micro device fabrication will be described. Here the mechanical micro tooling and the local metal deposition control are to be combined. The small robot simply consists of piezoelements for source of motion and electromagnet for machine clamping so that it can provide stable locomotion with sub-micron resolution on any curved surface even in the high vacuum chamber[1]. Unique small tools which can be remotely controlled are mounted on the robots so that they can collaborate in doing microscopic operations[2],[3]

In the preliminary experiments, the local deposition patterning control of metals is carried out by two miniature robots. The precise deposition pattern is obtained by both the accurate positioning action of two miniature robots in lateral and the thickness control of deposition time. Then another miniature robot with a small manipulator is also ready for supplying different material grains into the

heated pot. So this simple sequence can provide a set of unique micro deposition pattern with difference materials. Furthermore another small robot with a hopping micro tool of diamond is developed for the micro indentation on the specified sample. The demonstrative results such as a batch of micro devices for the humidity sensor and micro pit holes are successfully achieved by the collaboration of miniature robots.

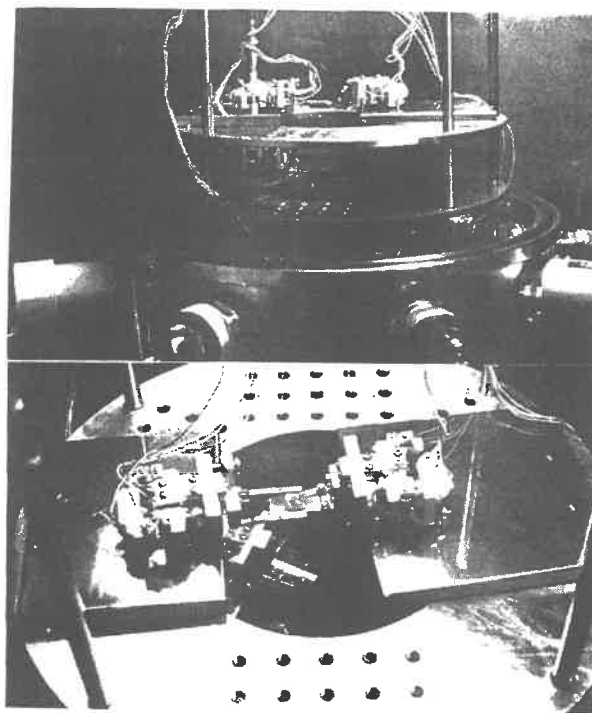


Fig.1 Experimental set-up
(3 miniature robots with mask, substrate
and manipulator in the deposition chamber)

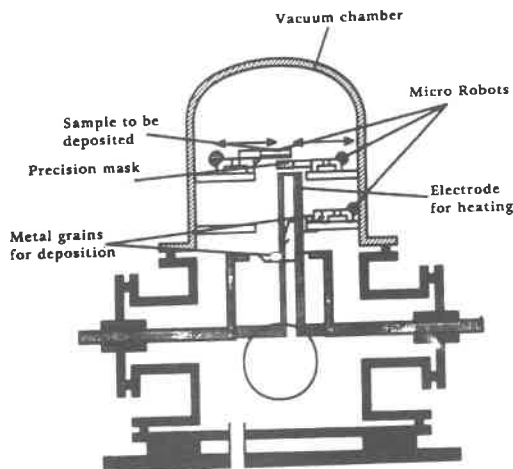


Fig.2 Layout of the experimental facility

2. Collaborative Work of Miniature Robots for Flexible Micro Patterning.

Figure 1 shows the experimental set-up that three mini-robots with a mask, a substrate and a small manipulator can work to control the local deposition process in the vacuum chamber. Each small robot is required to work there and controlled remotely by the responsible computer. There are two stages in the chamber, the upper stage for two small robots which are responsible for controlling the positions of the mask and the substrate precisely while the lower one is for the robot which will supply the metal grains into the heat pot as illustrated in Fig.2. At first, the mask and the substrate can be overlaid and positioned at the initial point by two micro robots' collaboration so that they can move incrementally with sub-micron resolution. At the bottom of the stage center, there is the heating pot for metal deposition and evaporated metal can go through the mask to the substrate. This arrangement provides local deposition patterns at any location of the substrate easily although the uncertainty of the pattern edge is determined by the layout of the deposition source, the mask and the substrate. On the lower stage, another mini-robot stands by for supplying different metal grains, for example Cu, Au and Al into the heat pot on demand so that the small amount of metal can be evaporated quickly. This process of small three robots action can be repeated to make the micro patterns anywhere as illustrated in Fig.3.

Experimental results to indicate the basic

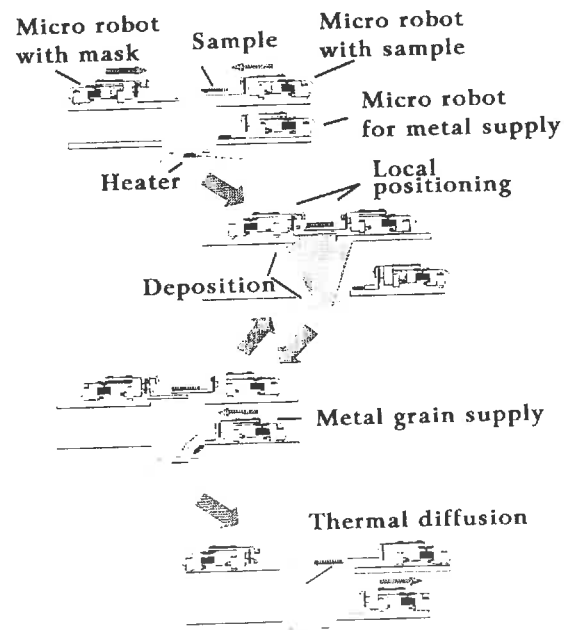
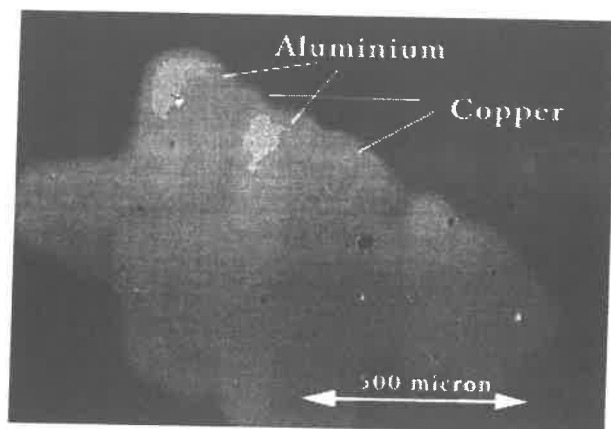


Fig.3 Process of flexible local deposition by miniature robots

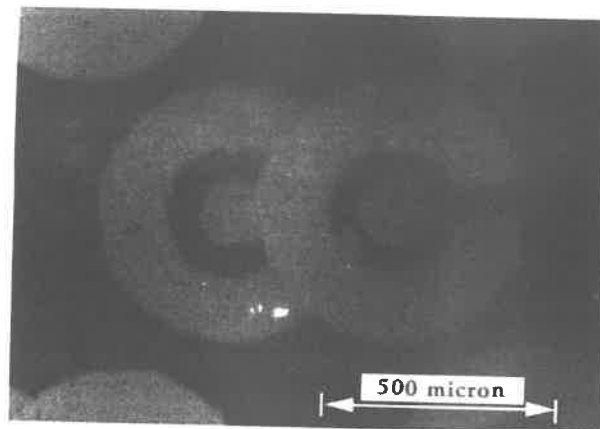
performance are shown in Fig.4. The set of the patterns of crossed arrow feature which are made of copper and aluminum have been deposited on the glass substrate with the same separation of 100 micron alternately in Fig.4(a). The unique pattern which is possibly applied to the humidity sensor can be also patterned with the simple control sequence of micro robots as shown in Fig.4(b). In the primary experiments, all of micro robots are maneuvered manually by the operators and the open loop positioning control manner based on the count of steps is used although the automatic production system are currently developed.

3. Hopping Tool for Micro Indentation

In the conventional IC batch process, the photo lithography and etching techniques are widely used for material removal and patterning. We have been considered the possibility to combine the fine mechanical machining and the physical/chemical surface processing for the local micro fabrication. Here we may propose the hopping tool which is excited at the resonance frequency on the small robot and synchronized with the robot locomotion for dynamic micro-indenting to the thin film sample. The hopping tool consists of the micro tool



(a) Cross-arrow patterns of aluminum and copper



(b) Speical pattern for a humidity sensor

Fig.4 Experimental results made by the miniature robots

attached at the end of a cantilever and the piezo element as a resonator as illustrated in Fig.5. The distance between the tool end and the surface is adjusted by the micro screw so that the tool end can slightly contact to the surface. The mechanical frame of the tool hopping which is excited by the piezo is designed to resonate for expanding the tool stroke at the same frequency of the robot locomotion like an inchworm. This arrangement can make the dynamic micro indenting along the robot's path with the accurate stepwidth. The end of tool may be changeable for several kind of feature such as conical, quadrangular pyramid and hemisphere forms which are to be also ultra precisely machined as shown in Fig.6.

The typical experimental result is shown in Fig.7(a), the photograph by an optical microscope, that is a set of indented pits by the conical diamond tool on the aluminum plate. We succeeded in making precise micro indentions

with the separation of 10 micron. Another sample of SiO₂ sputtered Si substrate is also used as the target of hopping indentation as shown in Fig.7(b). Smaller pits of approximately 1 micron diameter can be achieved at every 10 micro separation. It is obvious that if greater indentation force will be applied then the surface become cracked due to the nature of brittle material. However such tiny indenting force in this experiment is enough small to keep the ductile mode machining. It is expected that the thickness of SiO₂ thin film on Si is approximately ~200 angstrom. So the end of pit may be extended to the Si bulk although

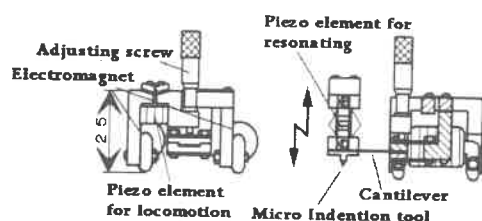


Fig.5 Micro robot with a hopping tool for micro indentation

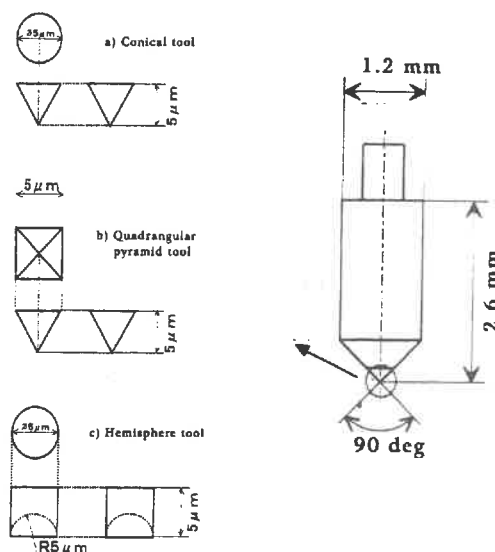
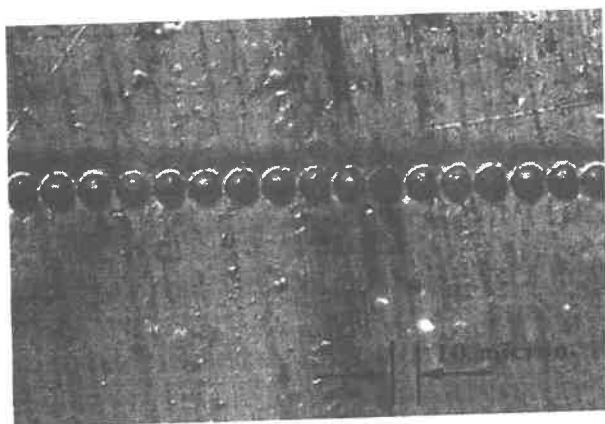
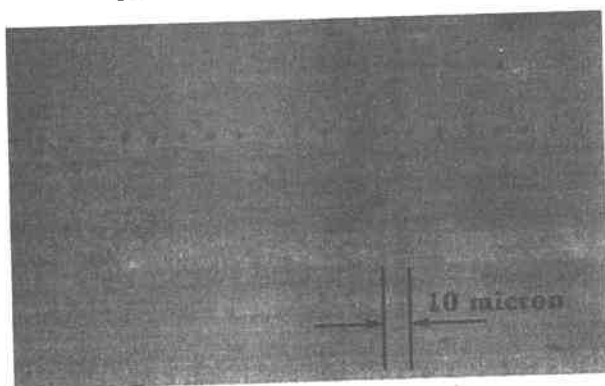


Fig.6 Micro tool top exchangeable for different indentation.



(a) Accurate indentions by the hopping tool of the robot on aluminum substrate



(b) Micro indentions on SiO₂ sputtered Si substrate

Fig.7 Dynamic micro indentions performed by the micro robot

we need much more detail investigation. It is feasible that some semiconductor devices can be produced by using this unique process performed by micro robots as shown in the conception of Fig 8. With the help of the thermal diffusion effect, it might be possible to get doped material around the indented hole so that the functional effect of such MOS-FET device can be produced.

4. Conclusion and Future Works

We proposed the new fabrication system which was the combination of the flexible micro deposition patterning as a physical process and the micro indentation by the hopping tool of micro robot as a mechanical machining. Some experimental results such as a set of deposition metal and micro indentation to the SiO₂ thin film were demonstrated. This micro robots system surely provides the potential application to a

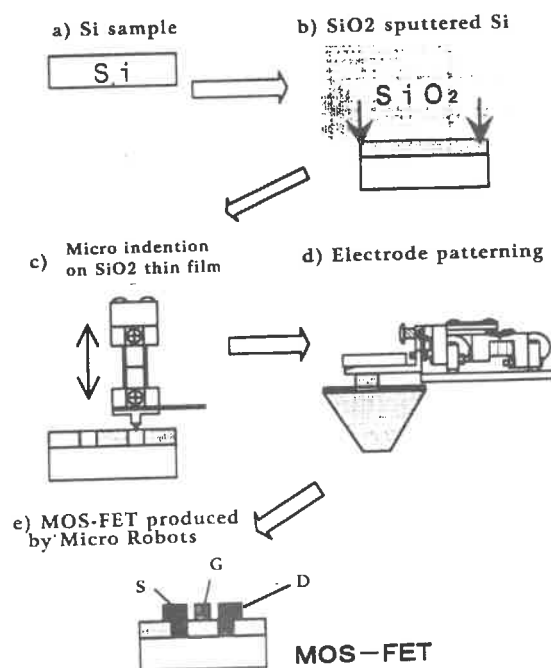


Fig.8 Micro fabrication process to make semiconductor devices by the micro robots

flexible micro device fabrication with cost saving benefit and low energy consumption. Currently we are on going development to establish "Desktop Micro Robots Factory" which can produce micro devices, assemble micro parts and could be installed into the conventional machine systems. Parts of this research were sponsored by Tateishi Science and Technology Fundation.

References

- [1] H.Aoyama, K.Mizutani, F.Iwata and A.Sasaki; Micro-Pattern Control by Miniature Robots in Vapor Deposition Process, *Proc. of 10th ASPE Meeting*, (1995)pp.384-387
- [2] H.Aoyama, F.Iwata and A.Sasaki; Collaboration Precise Operation by Two Miniature Robots - Single Diamond Cutting with Local Position Feedback, *Proc. of 9th ASPE Meeting* (1994)pp.238-241
- [3] H.Aoyama, F.Iwata and A.Sasaki; Desktop Flexible Manufacturing System by Movable Miniature Robots-Miniature Robots with Micro Tool and Sensor, *Proc. of IEEE Int. Conf. on Robotics and Automation*, Vol.1(1995)pp.660-665

Authors Index

Ahn, Hyeong-Joon	406
Ahn, Jung-Hwan	215
Ahn, Woo Jung	434
Alayli, Y.	173
Aoyama, Hisayuki	529
Asamoto, Kohsaku	466
Ascanio, Gabriel	139, 283, 369
Atherton, Paul D.	328
Aziz, David J.	157
Badami, Vivek G.	153
Badrawy, S.	49
Barylski, Adam W.	287
Bernhard, David	49
Bhattacharya, R. N.	143
Bifano, Thomas G.	515
Birkicht, Albert	**
Blaedel, Kenneth L.	271
Bonis, Marc	173
Borchardt, Bruce R.	443
Bosse, H.	131
Bourdet, P.	82
Brinksmeier, Ekkard	33
Brooks, Charles W.	410
Brown, Christopher A.	118
Bruning, John H.	25
Caballero, Alberto	139
Cava, Carlos J.	283
Chang, C. M.	**
Chang, In-Bae	406
Charlton, Thomas	447
Chen, Jenq Shyong	422
Chicurel, Ricardo	283
Cho, Sung Wook	434
Chou, Y. K.	451*
Chu, Xiohna	**
Chun, Keyoung Jin	345, 351
Coane, Philip	458
Copeland, Nathan R.	410
Corbett, John	365
Coyne, Dennis	15
Cuttino, James F.	390, 394, 398
Daito, Masamichi	211
Daniel, Cecil M.	65
Davies, Matthew A.	451*
Davis, Pete J.	271
Devitt, Andrew J.	377
Di Giacomo, Benedito	53
Díaz, Rufino	139
Dickmann, K.	511
Ding, Yao	78
Dörner, Peggy	319

*Extended abstract not available at press time.

Dow, Thomas A.	230
Drescher, Joseph D.	61
Duduch, Jaime G.	224
Dutterer, Brian S.	410
Dyer, Stephen W.	502*
Eberhardt, Keith R.	443
Edgeworth, Robert	90, 386
Estler, W. Tyler	29, 443
Evans, Christopher J.	29, 451*
Fähnle, Oliver W.	402
Fang, Fengzhou	220
Fawcett, Helen E.	515
Feautrier, Dominique	82
Fleig, Jon F.	25, **
Frankena, H. J.	402
Frase, G.	131
Freudenberg, Heiko	507
Friedrich, Craig R.	458
Fu, Jiyou	78
Fujita, S.	211
Fukada, Shigeo	333
Furutani, Katsushi	307, 357, 466
Galindo, Enrique F.	369
Gao, Wei	57, 414
Gatzen, Hans H.	454
Gee, Anthony E.	365
Geerkens, J.	490
Gim, Taeweon	365
Goch, Gert	490
Goettert, Jost	458
Goho, Kenji	110
Gopinathin, Niranjana	458
Greene, Wade	187
Griffith, Lee	271
Gweon, Dae-Gab	337, 341, 478, **
Ha, Jae Y.	195
Hamaker, Chris	178
Han, Dong-Chul	351, 406
Hara, Seiichiro	122
Hashiguchi, Makoto	191
Hashimoto, Hiroshi	262
Hatamura, Yotaro	238, 242, 482, 521
Hatsuzawa, Takeshi	41
Hayashi, Akira	110
Henrio, J.C.	173
Hicks, Thomas R.	328
Hilbing, Robert	498
Hirayama, Takao	311
Hoch, Pierre	118
Hocken, Robert J.	386, 474
Holmes, Michael L.	386, 474

**Abstract unavailable at press time.

Holzhauser, Wolfgang	418
Hopp, T.	443
Horaguchi, Iwao	258
Horio, Ken-ichiro	199, 275
Hosokawa, Akira	203
Houston, J. E.	19*
Huang, Chin-Chou	361
Hümmeler, Joachim	486
Imai, Kenichiro	262
Inaba, Fumio	211
Ishii, Kazuhisa	482
Itoh, Naoki	357
Iwata, Futoshi	529
Jaraus, Konrad	19*
Jared, Bradley H.	230
Jasinevicius, Renato G.	224
Jersch, J.	511
Joe, Yeo-Uk	165
Joung, Yoon	406
Jung, H.	**
Jusko, Otto E.	131
Kaiser, Mark J.	69
Kakumanu, Thrinath S.	250
Kallenbach, Eberhard K.	507
Kamiya, Kazuhide	169
Kanai, Akira	211
Kanai, Masaya	414
Kasai, Toshio	199, 275
Kato, Ryoji	211
Kato, Shigeo	315
Kim, Hyung Min	86
Kim, Oui-Serg	406
Kim, S. M.	337
Kim, Seung-Woo	426
Kim, Tae H.	195
Kim, Won-jong	470
Kiriya, Zenkichi	191
Kiyono, Satoshi	57, 414
Klein, Martin	21
Klein, Stephan	21
Knapp, Byron R.	246
Knight, P. Douglas, Jr.	106
Kohno, Tsuguo	183
Kondo, Sumihisa	203
Koteswara Rao, K. G. R.	299
Koyabu, Kunio	279
Kramer, Martin	157
Kroll, Jeremy	37
Kuhlbusch, W.	507
Kulawiec, Andrew W.	25
Kuppuswamy, G.	299
Lau, Kam C.	**
Lawton, Kevin M.	149
Lee, Byung-Chai	86
Lee, G.	114

Lee, Ho-Jae	426
Lee, Jae Y.	195
Lee, Jin-Koo	165
Lee, Shuo-Jen	**
Lee, Sung-Q	341, 478
Lee, Taik Min	195
Leenders, M.H.A.	503
Levenson, M. S.	443
Li, Junqi	191
Li, Nan	37
Li, Zhen	73
Lim, HanSeok	215
Lim, Seong-Eun	426
Lin, Enhao	135
Liu, Shugui	73, 102
Lovingood, John	386
Lu, Y. S.	262
Lückel, Joachim	507
Lüdicke, Frank	131
Ma, Evan	143
Ma, Quanhe	**
Ma, Xin-hui	102
Maekawa, Katsuhiko	291
Mätzig, J. Christoph	454
Magara, Takuji	307
Mahdi, Mofid	303
Maißer, Peter	507
Malkin, Gene	430
Marinescu, Ioan D.	234
Marsh, Eric R.	207, 246, 295
Martin, Osborne A.	37
Martinez Orrego, Roxana	53
Martyniuk, Jerry	178
Matsumoto, Kiyoshi	242
Matsuo, Tetsuo	234
Matsuse, Atsuhiko	275
McClain, M.	443
McClimans, Donald P.	**
McConnell, Malachy	328
Melvin, B.	443
Miller, Michele H.	250, 254
Mirro, Gene	178
Mishina, Satoshi	161
Mitsui, Kimiyuki	110
Mitsuishi, Mamoru	191
Miura, Tadahiro	529
Miyashiro, Hiroshi	169
Miyashita, Masakazu	211
Miyazawa, Kimio	161
Mizuhara, Kiyoshi	381
Mizutani, Ken'ichiro	529
Mohri, Naotake	307, 357, 466
Moon, Kee S.	126, 98
Moon, Sungho	94
Mou, Jong-I	114

Mu, Yunjun	135
Muñiz, Abdul	139
Murakami, Kenji	529
Muralikrishnan, Balasubramaniam	386
Na, Byung Chul	351
Nagalla, K.	37
Nagao, Takaaki	191
Nakao, Masayuki	238, 242, 482, 521
Namperumal, Ramesh	386
Natarajan, A.	143
Nava, Rigoberto	283
Nickles, C.	447
Nijisse, Gert-Jan P.	373, 503
Nomura, Takashi	169
O'Reilly, Thomas B.	295
Ogura, Ichiro	414
Ohira, Fumikazu	279
Ohno, Naohiro	183
Ohta, Kazushige	238
Okada, Akira	525
Okamoto, Yasuhiro	525
Okuda, Seiichi	169
Olson, Walter W.	65
Ozawa, Norimitsu	381
Pahk, Heui-Jae	434
Palomino, David M.	369
Pan, Y. L.	**
Parks, Robert E.	29
Patterson, Steven R.	37, 106, 149, 153
Pereira, Paulo E.	386
Perkins, Jeffrey N.	**
Petersen, Bernd	486
Phillips, Steven D.	443
Piscotty, Mark A.	271
Preuß, Werner	33
Pril, Wouter O.	438
Qu, Weimin	254
Radice, Josh	377
Reigl, M.	490
Reséndiz, Roberto U.	283
Rhead, P.	328
Rhorer, Janelle L.	452
Rhorer, Richard L.	410, 452
Ritter, R.	490
Robles, Guillermo	369
Ruhbach, Lars P.	454
Ruiz, Leopoldo	139
Russell, Phillip E.	19*
Saeki, Morihiko	203
Saffert, Eugen	507
Saito, Theodore T.	271
Sakai, Katsuhiko	311
Sanders, Benjamin	45
Sasajima, Kazuyuki	122
Sato, Yutaka	258

Sawyer, D.	443
Schäffel, Christoph	507
Schaut, Adam J.	207
Schellekens, Piet H. J.	438
Schinstock, Dale E.	187, 390
Schmidt, David M.	37
Schmitz, B.	490
Schmitz, Tony	49
Sekiguchi, Kazuhiro	199
Sethupathi, B.	37
Shao, L-Z.	29
Sharma, Aditya	386
Shen, Yin-Lin	94, 443
Shibatani, Kazuhiro	357
Shim, Jong Y.	**
Shin, Dae Young	345
Shintani, Tadashi	315
Sotiropoulos, Nick	377
Spronck, Jo W.	373, 503
Sprongl, P.	490
Struik, Klaas G.	438
Subrahmanyam, P.	494
Sullivan, Paul J.	29
Sun, Z. N.	262
Sutherland, John W.	65
Takashima, Yasuhiro	238
Tanaka, H.	462
Tanaka, Shuji	521
Tang, W.	131
Tashiro, Hatsuzo	169
Thomas, Timothy N.	178
Todd, Michael W.	390, 398
Trumper, David L.	470, 474, 494
Tsukada, Tadao	122
Uda, Yutaka	183
Uno, Yoshiyuki	525
van Beek, H. Frank	503
van Brug, H.	402
van der Laan, C.J.	402
van Putten, P. F. A. M.	373
Venkatesh, Vellore C.	220
Vieira Porto, Arthur J.	224
Vieira Sato, Denise P.	53
Wan, De-an	135
Wang, Danping	173
Wang, Jianli	73
Wang, Wanjun	143
Wang, Yiding	98, 126
Ward, David E.	443
Weck, Manfred	486, 498
Weikert, Sascha	323
Wilhelm, Robert G.	37, 90
Williams, Mark	494
Wroblewski, Donald	515
Wu, Yan	394

Xu, Ying	328	Yazawa, Takanori	183
Yamada, Hisashi	307	Yin, Ling	234
Yamada, Yoshiaki	525	Yokotsuka, Tadashi	211
Yamaguchi, Katsumi	258, 311	Yoshida, Y.	262
Yamanaka, Takashi	311	Yoshikawa, Kazuo	169
Yanagi, Kazuhisa	161	Zarudi, Irena	267, 462
Yang, Fang	102	Zhang, Guoxiong	73, 78
Yang, Min Yang	195	Zhang, Liangchi	267, 303, 462
Yang, Yun-Won	165	Ziegert, John C.	45, 361
Yasui, Heiji	203	Zirn, Oliver	323